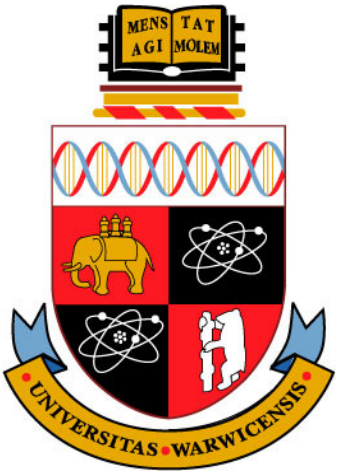
Investigations into the Isolation, Structure Elucidation and Biosynthesis of Bioactive Natural Products

Thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in Chemistry



Mansoor Ahmad
(ID: 0531314)

Supervisor: Professor Gregory L Challis

October 2011

Department of Chemistry
University of Warwick

Investigations into the Isolation, Structure Elucidation and Biosynthesis of Bioactive Natural Products

Part I: Investigation of fish-toxins produced by *Streptomyces* species

Part II: Structural and Biosynthetic Investigations of Foroxymithine

TABLE OF CONTENTS

Part II: Structural and Biosynthetic Investigations of Foroxymithine

Acknowledgements

In the name of Allah, The One, Most Gracious, Ever Merciful

I would like to dedicate this work to my parents. I believe that I would not have managed to get to this stage today without their constant support and continuous prayers.

I would like to take this opportunity to say that how very privileged I feel, to be a part of Challis Group. It has been a blessing for me to have a supervisor like Greg. It has been a long journey for me and I would like to give heartful thanks to Greg for being very patient and extremely understanding and accommodating of my personal circumstances during my time at Warwick. Greg is a brilliant scientist and has always inspired me with his extremely sharp intellect and vast knowledge of the subject and photography skills.

There is another set of three people I want to say special thanks to who have been through this journey with me all the way; my wife Asfa and two beautiful sons, Fahad and Adeeb. Their constant support and patience helped me to go through this time. I would also like to acknowledge the support of my sister, Kirn in Pakistan.

I cannot thank Lijiang enough for his help and advice starting from the very first day I came to Warwick. He has been extremely helpful over the years. To Christophe for his invaluable suggestions, optimism throughout and help in proof-reading my thesis. To Sarah for her help and advice on proof-reading. To Amaël for his help in synthesis work. To Ivan for useful advice and carrying out NMR experiments. To our lab technician Anne for her support. Last but not least thanks goes to a whole bunch of great people in the lab, past and present, including Malek (a very polite individual and great company), Nadia (remarkable support), Daniel, Stuart, Prakash Lauren, Lona, Joana, David and Nichola.

Finally a big thanks goes to the Environment Agency and the EPSRC for providing financial support for this work.

Declaration

The experimental work reported in this thesis is original research carried out by the author, unless otherwise stated, in the Department of Chemistry, University of Warwick, between September 2006 and October 2010. No material contained herein has been submitted for any other degree, or at any other institution.

Results from other authors are referenced in the usual manner throughout the text.

Mansoor Ahmad

Date:

Abstract

This project is divided into two parts. The first part of the project involved the investigation of fish toxins produced by *Streptomyces* species. Samples for fish-exposure experiments were prepared from an actinomycete belonging to the *S. griseus* clade isolated from a site of major fish kills, as well as from the type strain, *Streptomyces griseus* DSM 40236. Fish-exposure samples were prepared through a series of consecutive HPLC separations. We were able to narrow down the highest level of fish toxicity to four fractions, each containing only a handful of compounds. Comparison of the metabolic profile of the most toxic fractions showed that many of the compounds were common to all of them e.g. compounds yielding ions with *m/z* = 213, 241, 239, 301 and 309 were present in all the fractions. Some of these compounds were purified and analysed by high resolution mass spectrometry to determine their molecular formulae. A sample preparation and purification protocol for the fish toxins has been developed in this work. It was shown that the *S. griseus* type strain produces ichthyotoxic metabolites in addition to the environmental strain. This is a novel and unexpected observation.

The second part of the project involved structural and biosynthetic investigations of the iron-chelator and natural Angiotensin Converting Enzyme (ACE) inhibitor foroxymithine. The high structural similarity of foroxymithine to the known siderophore coelichelin and the lack of unambiguous experimental evidence in the literature to support the proposed structure of foroxymithine provided the impetus for these investigations. Foroxymithine was purified from a novel source, the culture supernatant of *Streptomyces narbonensis*. The gallium complex of purified foroxymithine was prepared and analysed by one- and two-dimensional high-field NMR experiments (^{1}H, COSY, HSQC, HMBC, NOESY, TOCSY and Difference NOE), which allowed the ^{1}H and ^{13}C NMR signals for the complex to be assigned.

Careful inspection of the ^{1}H NMR spectrum of Ga-foroxymithine revealed two signals (a major and a minor) for several of the protons. The origin of these signals was investigated using variable temperature and ^{71}Ga NMR experiments, and by LC-MS

analyses on a homochiral stationary phase. The duplicate signals were proposed to be associated with the existence of conformational isomers of Ga-foroxymithine.

The absolute stereochemistry of the four amino acid constituents of Ga-foroxymithine was determined using Marfey's method. Authentic standards of two of the anticipated acid-hydrolysis products of Ga-foroxymithine, D- and L-N5-hydroxyornithine were chemically synthesized to facilitate the Marfey's analysis. The results confirmed that foroxymithine contains L-N5-hydroxyornithine. Similar analysis was performed using the authentic standards of D- and L-serine and the results confirmed that foroxymithine contains L-serine. Marfey's method was also applied to elucidate the absolute stereochemistry (previously unknown) of coelichelin. Fe-coelichelin was purified from *Streptomyces coelicolor* M145. Marfey's derivatised coelichelin hydrolysate and Marfey's derivatives of L-Thr, L-allo-Thr, D-allo-Thr were analyed. The results showed that coelichelin contains D-allo-Thr. Similarly, analyses were carried out using the Marfey's derivatives of chemically synthesised D- and L-N5-hydroxyornithine, however the results were inconclusive.

The biosynthesis of foroxymithine in *S. narbonensis* was also investigated. A gene fragment proposed to be within the putative foroxymithine biosynthetic gene cluster was amplified by PCR from the genomic DNA of *S. narbonensis*. The gene fragment was cloned into a plasmid vector and sequenced. This confirmed that it encodes part of a putative formyl transferase that could be involved in foroxymithine biosynthesis. Fosmid libraries of *S. narbonensis* genomic DNA were prepared. Despite exhaustive efforts, a fosmid clone containing the putative formyl transferase encoding gene fragment could not be identified via PCR based screening of the library.

Abbreviations

A	Adenylation (domain)
aa	amino acid
ACE	Angiotencin converting enzyme
ACP	Acyl Carrier Protein
ADP	Adenosine diphosphate
ahfOrn	Acetylated hydroxyformylornithine
Ala	Alanine
amp	Ampicillin
Asp	Aspartic acid
AT	Acyl Transferase (domain)
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
C	Condensation (domain)
^{o}C	Celsius
CD	Circular Dichroism
CDA	Calcium-dependent antibiotic
CoA	Coenzyme A
COSY	Correlation spectroscopy
2D	Two dimensional
DCM	Dichloromethane
DH	Dehydratase (domain)
DMSO	Dimethylsulfoxide
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
E	Epimerisation (domain)
EDTA	Ethylenediaminetetraacetic acid
EIC	Extracted ion chromatogram
ER	Enoyl Reductase (domain)
ESI-MS	Electrospray ionization – Mass spectrometry
EtOAc	ethyl acetate
EtOH	Ethanol
FAD	Flavin Adenine Dinucleotide
FAS	Fatty Acid Synthase
FMN	Flavin mononucleotide
g	gram
Ga-fm	Gallium-foroxymithine
GC	Guanine-Cytosine
hfOrn	Hydroxyformylornithine
Glu	Glutamic acid
hOrn	Hydroxyornithine
His	Histidine
HMBC	Heteronuclear Multiple Bond Correlation
HPLC	High-performance liquid chromatography
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear Single-Quantum Correlation

hyg	Hygromycin
kan	Kanamycin
kb	kilo base pairs
KR	Ketoreductase (domain)
KS	Ketosynthase (domain)
LB	Luria-Bertani (Medium)
LC	(High Pressure) Liquid Chromatography
LC-MS	Liquid Chromatography – Mass Spectrometry
M	Methylation (domain)
Me	Methyl
MeOH	Methanol
mg	Milligram
MHz	Megahertz
MIB	Methyl isoborneol
min.	Minute
mL	Millilitre
mM	Millimolar
mm	Millimetre
mRNA	Messanger RNA
MS	Mass Spectroscopy
m/z	Mass to charge ratio
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear overhauser effect spectroscopy
NRPS	Nonribosomal peptide synthetase
nt	Nucleotide
OD	Optical Density
ORF	Open Reading Frame
Orn	Ornithine
PCR	Polymerase Chain Reaction
PCP	Peptidyl Carrier Protein
PKS	Polyketide synthase
PPTase	Phosphopantethienyl transferaze
Pro	Proline
^{R}	resistante
RBS	Ribosome binding site
RNA	Ribonucleic acid
rRNA	ribosomal RNA
rpm	revolutions per minute
^{S}	sensitive
SAM	S-adenosylmethionine
CDS	CoDing Sequence
SEM	Scanning electron microscopy
Ser	Serine
SFM	Soya Flour Mannitol medium
Sfp	Surfactin phosphopantetheinyl transferase
SMM	Supplemented Minimal Medium
T	Thiolation (domain)
Taq	Thermus *aquaticus* (polymerase)

TBE	Tris-boric acid EDTA buffer
t-Bu	*t*-Butyl
TE	Thioesterase (domain)
TEM	Transmission electron microscopy
Thr	Threonine
TLC	Thin layer chromatography
TOCSY	Total Correlation Spectroscopy
TOF	Time-of-flight
Tris	Tris(hydroxymethyl) aminomethane
tRNA	transfer RNA
UV	ultra-violet
VT	Variable temperature

List of Figures

List of Tables

Part I: Investigation of fish-toxins produced by *Streptomyces* species

CHAPTER 1 - INTRODUCTION

1.1 The impact of environmental pollution on fish

Water pollutants are a frequent problem for freshwater fish. These pollutants come from homes, industrial waste, mining etc. Fish gills are responsible for respiration, acid-base regulation and excretion of nitrogenous waste products and they are the first target of ichthyo toxic water pollutants. This is why fish gills are a very good indicator of water quality and may also be used a model system to study ion transport across cellular and epithelial membranes in toxicant-affected environments [1-3]. Many studies have shown that damage to the gill epithelium results in hypoxia, which is caused by inhibition of gas exchange at the gills [4-6].

Heavy metals like Cu^{2+} and Hg^{2+} have been shown to have severe effects (both morphological and physiological) on fish gills. Physiological effects involve a reduction in ion levels in blood or acidosis, and inhibition of enzymes such as Na^+, K^+-activated ATPases and carbonic anhydrases which play important roles in ion transport [7]. Low pH also has severe effects on fish gill morphology. When brook trout was exposed to pH 5.2 for 7 days, separation of epithelial layers of the secondary lamellae on gills was observed [8]. Similar effects on the gills of rainbow trout were observed when exposed to the anionic detergent sodium lauryl sulphate [9]. In another study, exposure of Yearling Sunapee trout to pH 4.5 for 8 days showed slight swelling of secondary lamellae and lowering the pH to 4.0 made this worse [10]. Other effects of the exposure of fish gills to acidic pHs, include a significant change in the phospholipid composition of gill tissue accompanied by a substantial increase in the percentage composition of unsaturated fatty acids [11, 12].

The toxic impact on the gills of air-breathing catfish *Heteropneustes fossilis* of lethal concentrations of lead nitrate has been studied by Ram Sanehi and Tarun (2002). Rapid change in the gill morphology was observed which involved detachment and lifting of the epithelial linings from the surfaces of the primary and secondary lamellae of fish gills. At the same time uncontrolled growth of primary and secondary lamellae took place and consequently gills appeared to be a solid mass of cells with

almost no free surface left for gaseous exchange. Consequently blood capillaries of the gills collapsed, causing asphyxiation and the death of the fish [13].

1.1.1 Methods for identification of waters containing ichthyotoxic pollutants

It is becoming increasingly important to develop methods for rapid identification of waters polluted with ichthyotoxic substances, because the longer it takes to identify such waters, greater the chance of fish mortalities and widespread dispersion of the pollutants. Various techniques have been used in the past to investigate pollutant-affected fish gills. Light microscopy has been used but this does not give a clear picture of the extent of damage to the gills [14]. Other methods include applying morphometric techniques, which involve studying variation and change in the form (size and shape) of organisms, but these methods are complicated to apply [15, 16]. Recently electron microscopy-based techniques have gained popularity because the data obtained give much clearer pictures of pollutant-affected fish gills. One of these techniques is Transmission Electron Microscopy (TEM), which gives ultrastructural level information about the subject. The only disadvantage of this technique is that it is a bit time consuming and cannot be used as a tool for routine investigation. On the other hand Scanning Electron Microscopy (SEM) is fast and gives excellent quality results in investigations of the morphological changes to pollutant-affected fish gills [1, 2, 17]. This technique has been used routinely in recent studies. Lewis and Kirk used SEM to analyse changes in the gills of rainbow trout by exposing the fish to phenol, copper and ammonia (because these are the most abundant pollutants secreted into freshwaters in industrialized countries). Changes to the fish-gill lamellae were studied with increasing concentrations of pollutants. Use of SEM enabled larger areas of fish gills to be studied in a short time and this technique can be used routinely to monitor the quality of water environments [18].

1.2 The production of ichthyotoxic metabolites by microbes

There is continuing strong evidence suggesting that bacterial exotoxins are responsible for many unexplained and often large-scale fish mortalities in UK rivers, canals and lakes. There is also a link between bacterial isolates from affected or

'toxic' sites and significant pathological changes to the gills of fish. Previous histopathological studies of fish exposed to pollutants have shown that fish gills are efficient indicators of water quality. In an experiment where five fish species were exposed to cell cultures and filtrates of *Streptomyces griseus* strains, the fish showed some severe pathological changes. Their gills were found to display varying degrees of collapse, hyperplasia and fusion of secondary lamellae with loss of microridging, distortion and erosion of the filamental epithelium of primary lamellae, leading to the loss of respiratory surface, hypoxia and mortality. These effects worsened as the concentration of the cells was elevated (Figure 1-1). However, fish were found to recover from these conditions at low cell concentrations [19].

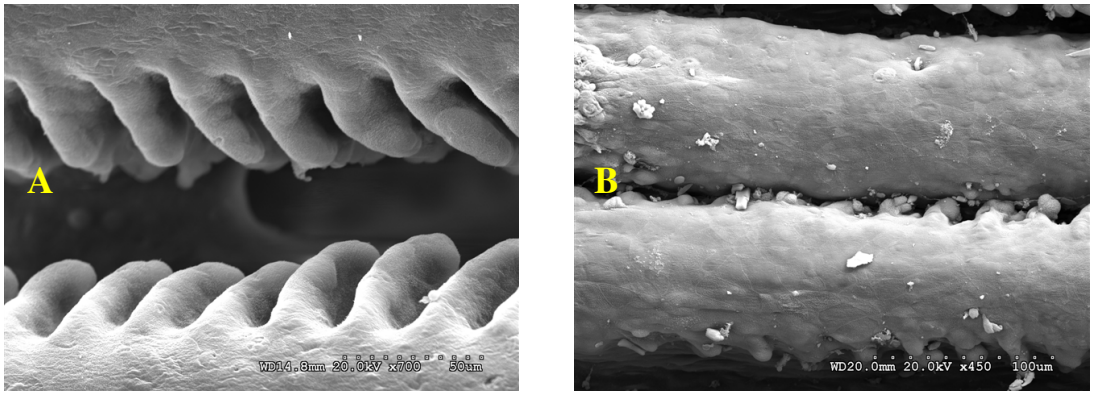


Figure 1-1 A-Control, to show equal spacing of secondary lamellae, B-Secondary lamellae of fish after exposure to filtrates of *Streptomyces griseus* showing complete fusion (Photos provided by Prof. John Lewis, School of Biological Sciences, Royal Holloway, University of London)

There have been many studies aimed at investigating fish mortalities due to bacterial toxins. Over 40 species of cyanobacteria are known to produce secondary metabolites that are toxic to marine animals. These toxins have been reported to cause the death of rainbow trout and salmon [20, 21]. In one case fish mortalities have been found to result from an algal bloom producing microcystin-LR in an Arkansas catfish production pond [22]. Ichthyotoxic metabolites are not essential in any known biochemical pathway and are accumulated in cells during the late exponential growth phase.

Microcystins are named after *Microcystis aeruginosa,* the cyanobacterium from which they were first isolated [23, 24]. These are a group of cyclic heptapeptide hepatotoxins produced by several cyanobacterial genera. The macrocycle consists of two proteinogenic amino acids and five non-proteinogenic amino acids. It is variation in

the two proteinogenic amino acids within the macrocycle that distinguish microcystin types from each other. The nonproteinogenic amino acids within the macrocycle are highly conserved. Each microcystin is named on the basis of the variable amino acids within its structure. For instance, one of the most common toxins found in water supplies around the world, microcystin-LR (Figure 1-2), contains the amino acids Leucine (L) and Arginine (R) in the variable positions [25].

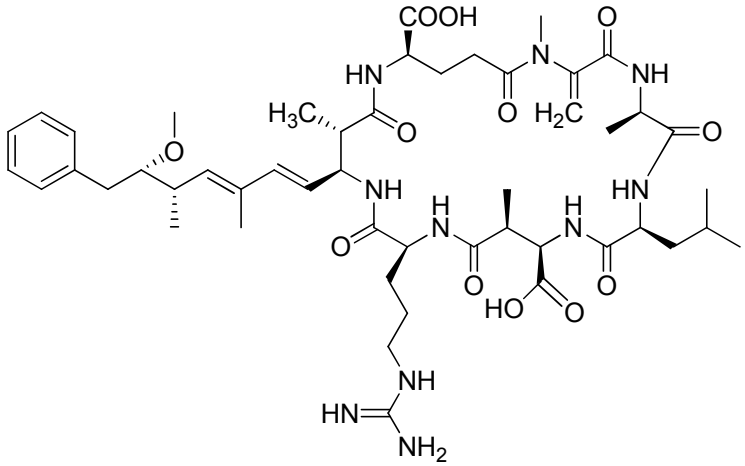


Figure 1-2 Structure of Microcystin-LR [26]

Microcystins are extremely stable in water. They can persist in both warm and cold water and can tolerate radical changes in pH. To date, approximately 75 different types of microcystin have been discovered. These hepatotoxins inhibit the protein phosphatases inside hepatocytes. As a result the hepatocytes collapse and cause lethal damage to the liver [26]. Microcystins cause Netpen Liver Disease (NLD) in salmon and rainbow trout [167]. Exposure to microcystin-contaminated water is responsible for illness in humans and causes sickness or even death in cattle and wild animals[25, 26].

1.3 Recent major episodes of fish mortality in fresh waters

A combination of rare, natural phenomena caused a major fish kill just outside Hungerford in West Berkshire, UK (Figure 1-3). More than 150 tons of fish were wiped out at the Berkshire Trout Farm and hundreds of thousands of coarse fish were killed in the Kennet and Avon Canal and the River Dun. The first major incident started on 4 March 1998. This recurred to a lesser degree in the spring of 1999.

Another similar incident took place in the Kennet and Avon Canal near Devizes in spring 2001. Yet another incidence of major fish mortalities took place in the Trout Lake in Hintlesham in 1999. Besides these major incidences, small scale fish mortalities resulting from unknown causes have been taking place in canals, lakes and rivers in recent years in the UK [19]. In most of the cases, fish deaths, were due to gill hyperplasia and the fusing of secondary gill lamellae.

Figure 1-3 Picture showing the large scale fish mortalities at a fish farm in Hungerford, West Berkshire, UK in spring 1998 (Photo provided by Ros Wright, Environment Agency)

An initial investigation after the first incident at Hungerford was carried out. Water contamination due to chemical spillage was ruled out because the levels of toxin continually fluctuated. It was found that only fish were affected, which rules out insecticides or pesticides as a cause, because these would have killed everything. Another line of investigation was based on the fact that, at that time, the unseasonably high temperature had caused an increase in algal blooms in the canal, and decomposing algae have been known to be associated with fish kills. Large quantities of polysaccharides were found in the affected waters, which are commonly associated with algae. It was suggested by the researchers that one of the causes of fish death was the over-accumulation of the polysaccharides, which can clog the fish-gills and suffocate them. This was consistent with the findings of pathological examinations of

the fish. Naturally-occurring bacteria, which can survive on polysaccharides and produce toxins that, if produced in high enough amounts may contribute to fish deaths was identified as another possible cause. As a quick fix, the affected water was treated with hydrogen peroxide which is an established plan of action for combating fish mortalities resulting from bacterial toxins. Laboratory trials were done on the affected water and it was found to be safe after treatment with hydrogen peroxide [27, 28].

Water and sediment samples from 17 out of 36 potentially toxic sites were analysed and, based on the results, toxic incidents were divided into three major groups: (a) actinomycetes-based, (b) cyanobacterial/algal based; and (c) Gram-negative bacteria-based. The majority of the incidents were actinomycetes-related. SEM studies showed fish-gills were different to those caused by toxins like heavy metals (copper, chromium and cadmium), cyanide, altrazine phenol or ammonia. This supports the hypothesis that naturally-occurring biological toxins are associated with the unexplained fish-kills. Initially a group of 11 actinomycete isolates were identified to be present at several of the sites of the fish mortalities. Six of these were identified on the bases of 16S rRNA sequencing. Out of the six five were found to belong to the *Streptomyces griseus* clade and one was found to belong to the *Streptomyces sampsonii* clade [19, 29].

The aim of this project was to take the project initially started by the Environment Agency to the next level and gain better understanding about the effects of a variety of metabolites produced by the fish-toxin producing strains of *Streptomyces griseus* on the fish and to be able to design a reliable bacterial metabolite sample preparation protocol.

1.4 General introduction to *Streptomyces*

Streptomyces bacteria belong to the order Actinomycetales [30]. Actinomycetes are unique non-motile, filamentous, Gram-positive bacteria [31, 32]. When they grow, they form branching filaments called mycelia, which are very similar in appearance to the mycelia of some fungi [31]. *Streptomyces* are the most important of the actinomycetes because they produce a large number of important antibiotics including streptomycin

[30, 33-35]. Over 500 species are known, a few of which are pathogenic to animals or plants [36, 37].

Streptomyces species are mainly found in soil. Much of the characteristic earthy smell of soils is due to geosmin and 2-methyl isoborneol given off by *Streptomyces* species [38-42]. *Streptomyces* can survive on almost anything including sugars, alcohols, amino acids, organic acids, and aromatic compounds [30, 43-47]. Their metabolic diversity is due to their extremely large genome which has hundreds of transcription factors that control gene expression, allowing them to respond to specific needs [48, 49]. *Streptomyces* species produce a large number of antibiotics and several other classes of biologically active secondary metabolites such as anticancer agents, immune suppressants, obesity drugs, parasiticides, insecticides and herbicides [30, 34, 50-56]. There are more than 5,000 antibiotics which have so far been isolated from *Streptomyces* spp [57]. Actinomycetes produce about two-thirds of the known antibiotics that are produced by microorganisms and amongst them nearly 80% are made by *Streptomyces* [30, 35].

1.5 *Streptomyces* growth

In addition to the similarities to fungi in their cellular structure, *Streptomyces* also resemble fungi in their elaborate life cycle (Figure 1-4) [30, 58, 59]. During the vegetative growth stage of *Streptomycete* development, DNA replication takes place without cellular division, creating a filamentous structure [46]. *Streptomycetes* reproduce and disperse through the formation of spores, after the period of vegetative growth [30, 60]. The spores are produced in aerial filaments called sporophores more commonly known as “aerial hyphae” [59], which rise above the colony [59, 61, 62]. Spores are a semi dormant stage in the life cycle of *Streptomyces* and can survive in soil for many years [46, 60]. Indeed, viable *Streptomyces* cultures have been obtained from 70 years old soil samples [63]. Because the complex life cycle of Streptomycetes resembles that of multicellular eukaryotes, researchers use *Streptomyces* as a model for the development of the more complex eukaryotic systems.

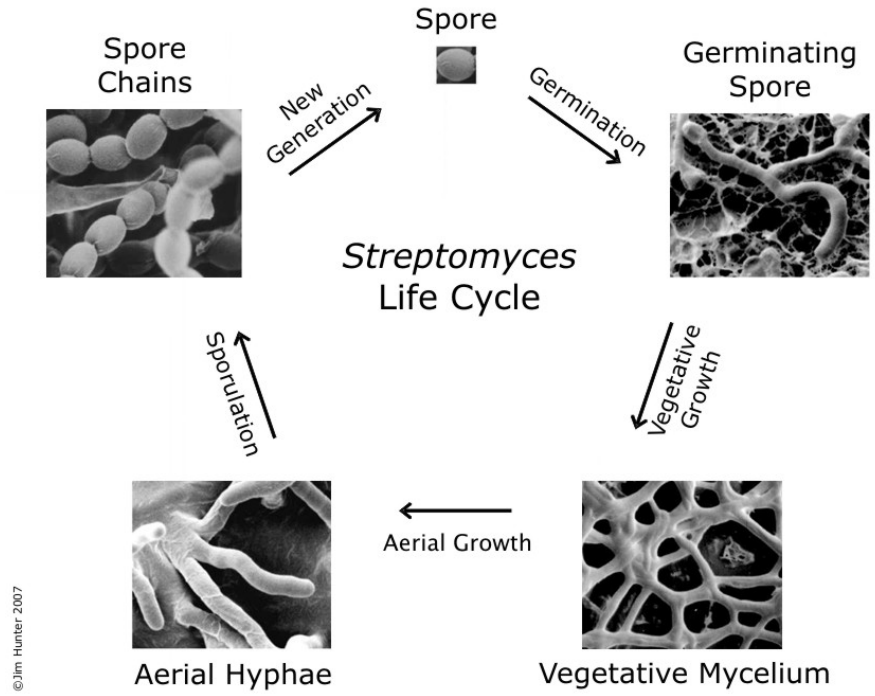


Figure 1-4 The life cycle of *Streptomyces coelicolor*[64]. A single spore germinates into vegetative mycelium, which is followed by aerial growth with the production of aerial hyphae. These hyphae in turn undergo synchronous septation to produce unigenomic spore compartments, which disperse and thus commence a new cycle

Streptomyces are usually considered to prefer neutral to alkaline environmental pH, although they have been found living in remarkably variable pH and nutritional conditions [44]. This is why Streptomycetes are the most common environmental isolates. They can grow in buildings at a wide variety of pH values in the presence of appropriate nutrients such as fungal mycelium, or water–damaged plasterboards supplemented with substrates including starch and covered with paper [44]. A relatively high number of *Streptomyces* species exist in soil as inactive spores for most of the time [65]. Spores can survive at very low levels of nutrients and water but the mycelial stage of growth is sensitive to low water levels [65]. Spore germination has also been found to be linked to population density within a species [66, 67]. It is suggested that some signalling factors are responsible for slowing down the germination process when the population density reaches a certain level [66, 68]. This might be a natural way of controlling cell density when nutrient availability is limited. In a study, aimed at assessing the diversity of aquatic actinomycetes in lakes within a Chinese province, Yunnan [69], it was found that actinomycetes are more abundant during the dry season than the rainy season.

Antibiotic production in Streptomycetes is generally growth phase-dependent. In the liquid media it begins as the culture enters stationary phase and in agar media it coincides with the beginning of morphological differentiation. Most antibiotics are produced through a complex biosynthetic pathway, with a cluster of genes (normally 20-30) specific to the synthesis of any one compound. These gene clusters usually contain pathway-specific regulatory genes which act as transcriptional activators, and which are themselves subject to control by pleiotropic regulatory genes [30].

1.6 *Streptomyces* genomes

The genome size of most *Streptomyces* species including *S. coelicolor* A3(2) [70], *S. lividans* 66 [71], *S. ambofaciens* [72], and *S. griseus* IFO03237 [73] has been estimated to be ~7.8-10.1Mb (Table 1-1). The genomes of at least four *Streptomyces* species have been fully sequenced. The entire genome sequence of *S. coelicolor* was published in May 2002 [74]. The linear chromosome is 8.7Mb long and is predicted to contain 7,825 genes, about twice as many as typical free-living bacteria such as the Gram +ve, non-spore forming and amino acid producing bacteria *Coynebacterium glutamicum* and *Brevibacterium lactofermentum* with the genome size of 3Mb and 2.9Mb, respectively [75, 76] and ~4.6Mb genome of a soil bacterium *Pseudomonas stutzeri* Strain DSM4166 [77]. The complete *S. avermitilis* genome sequence was published in 2003 [78]. It is 9,025,608 bp long, and contains 7,575 putative coding sequences (CDSs) [78]. The entire genome sequence of the streptomycin producing strain *Streptomyces griseus* IFO 13350 was published in 2008 [79]. The linear chromosome is 8,545,929 bp long, with an average G+C content of 72.2%, and is predicted to contain 7,138 CDSs. Most recently genome sequence of an insect associated *S. griseus* strain XylebKG-1 has also been published [80] with a total size of 8,727,768 bp. Table 1 shows some feature of the four *Streptomyces* strains whose genome has been sequenced including the two *S. griseus* strains.

Table 1.1 General features of the chromosomes of four sequenced *Streptomyces* strains; *S. griseus* IFO 13350 [79] , *S coelicolor* A3(2) [74], *S. avermitilis* [78], and *S. griseus* CylebKG-1[80]

Species	Length (bp)	Avg G + C content (%)	No. of protein coding genes	No. Of rRNA (16S-23S-5S) operons	No. of tRNA genes
S. griseus IFO 13350	8545929	72.2	7138	6	66
S. coelicolor A3(2)	8667507	72.1	7825	6	63
S. avermitilis	9025608	70.7	7583	6	68
S. griseus strain XylebKG-1	8727768	72.1	7265	6	66

Comparison of the published *Streptomyces* genome sequences reinforces the idea that *Streptomyces* species are highly adaptable enabling them to cope with a variable and complex environment. The *S. coelicolor* genome surprisingly contains 65 putative sigma factors (components of the bacterial RNA polymerase that play major roles in the transcription process) [81]. A staggering 45 of the putative sigma factors belong to the extracytoplasmic function (ECF) class, involved in the transcription of gene sets in response to various environmental stimuli [81]. This is not just the case for *S. coelicolor;* the number of putative sigma factors is almost the same in the *S. avermitilis* genome. Similarly, *S. coelicolor* has 18 genes encoding cytochrome P450 monooxygenases, presumably allowing the organism to cope with a range of complex xenobiotics in the soil [81].Some of these cytochrome P450 are likely to be involved in the biosynthesis of secondary metabolites [81]. Comparison of the of *S. griseus*, *S. coelicolor* and *S. avermitilis* genome sequences implies the existence of 24 *Streptomyces*-specific proteins [81].

In the context of biotechnology, the most interesting finding in the three published *Streptomyces* genome sequences is the abundance of genes that encode enzymes likely to be involved in secondary metabolism [74, 78, 79, 82-86]. Before the first genome was sequenced, gene clusters directing the biosynthesis of three antibiotics and a spore pigment were known to reside within *S. coelicolor*. The genome sequence revealed two dozen additional putative secondary metabolite biosynthetic gene clusters, likely to direct the production of pigments, complex lipids, signalling molecules, iron-scavenging siderophores and antibiotics [81]. In a preliminary report on the *S. avermitilis* genome prior to completion of its sequence, 25 putative secondary metabolite biosynthetic gene clusters were described [87]. This number increased to 30

upon completion of the sequence [81]. The *S. griseus* genome contains 34 gene clusters (or genes) that direct the biosynthesis of known or putative secondary metabolites. Remarkably, nearly all the gene clusters appear to direct the production of different compounds in the three species, indicating the large number of secondary metabolite biosynthetic remains to be discovered in the actinomycetes as a whole.

Streptomyces avermitilis has an important commercial and medicinal importance as the antibiotic avermectin is produced by this strain [88]. Since the publication of the whole genome sequence of this *Streptomyces* and others it has opened a whole new world for the discoveries of new secondary metabolites and potential useful drugs. The sequence analysis of *S. avermitilis* has revealed that there are 30 types of secondary metabolite gene clusters which corresponds to 6.6% of the genome and the products of most of these gene cluster are still to be discovered [89].

Similarly *S. coelicolor* genome revealed more than 20 gene clusters encoding for putative secondary metabolite [74] and based on the pioneer work based on gene cluster sequence analysis (also called genome mining) an important discovery of a secondary metabolite, a siderophore, was made by Challis *et al.* in 2005 [83]. The *S. griseus* genome was published in 2008 which contains 34 gene clusters or genes for the biosynthesis of known or unknown secondary metabolites [79]. Very recently sequence analysis of *S. griseus* has led to the identification of a peptide natural product called grisemycin [90].

1.7 *Streptomyces* as pathogens

In general *Streptomyces* have little pathogenic effect on humans but still there are reports where *Streptomyces* has been involved in human infections. A rare case of invasive infection causing pneumonia was reported to be caused by *Streptomyces* species in a patient with HIV infections [91] *S. somaliensis* has also been reported to be the primary cause of mycetomas in South America, Africa, India, Mexico, Malaysia and the United States [91-93]. There have been only very few reports of *Streptomyces* species causing any other type of infection other than mycetoma [91] but *Streptomyces*

species have also been described as the cause of septicaemia and some lung infections [94] and *S. griseus* has been reported to be responsible for a brain abscess [95].

Only a few diseases in plants are caused by these actinomycetes. The best known actinoymycete plant pathogen is *Streptomyces scabies* which causes common scab in potato (Figure 1-5) as well as in carrot, parsnip, turnip and other root crops [96]. Other species of *Streptomyces* including *Streptomyces setonii*, *Streptomyces. griseus*, *Streptomyces tendae*, and *Streptomyces aureofaciens* have also been associated with potato scab [96-98]. Potato scab is caused by plant toxin thaxtomin produced by *Streptomyces scabies* [96, 99]. Thaxtomin is a nitrated dipeptide phytotoxin which inhibits cellulose biosynthesis in expanding plant tissues, stimulates Ca^{2+} spiking, and causes cell death [100].

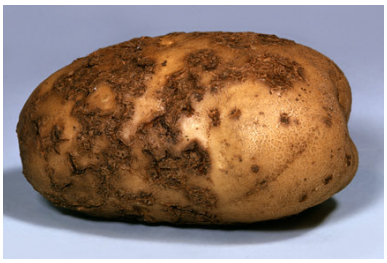


Figure 1-5 Potato scab caused by *Streptomyces* scabies

Health problems associated with water-damaged buildings have been linked to a variety of microbes and *Streptomyces* species have been suspected to cause health problems in such buildings by producing β-D-glucans, as well as cytotoxic and mitochondriotoxic metabolites [101-103]. *Streptomyces* along with other toxic bacterial species were identified in a dwelling where an occupant exhibited serious symptoms, including the exacerbation, asthma, sinusitis, urticaria, blocked nose, rhinitis, otitis, hoarseness, aching in joints, myalgia, and tiredness [104]. It was suggested that the presence of mitochondriotoxic actinobacteria like *Streptomyces* may adversely affect the tear film in the eyes and thus account for the eye irritation experienced by the inhabitant. Species of *Streptomyces* may also be involved in causing certain respiratory disorders in people living in mouldy houses [105, 106]. Andersson et al., have reported that extracts from water-damaged indoor building materials paralyze sperm and cause mitochondria to swell [101, 107]. They also found that *S. griseus* isolates obtained from dust and air in water-damaged buildings produce valinomycin and that paralysis of sperm and mitochondrial swelling occurred with pure cultures of *S. griseus* strains, as well as with commercially-obtained valinomycin.

1.8 *Streptomyces griseus*

In 1914 Krainsky was the first to report the isolation of *S. griseus* from Russian soil which was then called *Actinomyces griseus* [108]. Soon afterwards in 1916 Selman Waksman and one of his assistants isolated many similar cultures from American soil and carried on extensive studies on this microorganism [108, 109], This name was changed to *Streptomyces griseus* in 1943 [110].

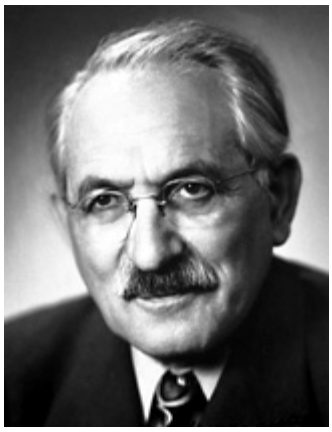


Figure 1-6 Professor Selman Waksman

The well-known antibiotic streptomycin is produced by one of the strains of this species and was initially reported by Albert Schatz in Waksman's laboratory in 1943 [111-113]. In recognition to his work on soil microbs Professor Selman Waksman (Figure 1-6) was awarded Nobel Prize of Physiology of Medicine in 1952 [110] . *S. griseus* is an ideal model to study prokaryotic development due to its ability to sporulate abundantly in liquid culture which is very uncommon for *Streptomyces* strains [20, 21, 58]. This characteristic of *S. griseus* has been exploited to study post-sporulation metabolic events initiated by phosphate or nitrogen starvation or nutritional downshift [20, 21]. Sporulation of *S. griseus* occurs at 26-28 hours of incubation while shaking at 30^{o}C in a defined medium. The timing of sporulation is unchanged even when the level of each nutrient is increased ten times. The levels of carbon and nitrogen sources, as well as phosphate, can be at a high levels when sporulation starts. This shows that in *S. griseus* sporulation is not associated with nutrient limitation [114].

1.9 Secondary metabolites of *Streptomyces griseus*

There is no one-line definition of the term secondary metabolite. These compounds usually have diverse, unusual and complex chemical structures [115]. They are not essential for the growth of the producing organism under laboratory conditions, but are likely critical for survival in the natural environment. They are usually produced after the phase of most active vegetative growth, when the organism is entering a

dormant or reproductive stage [35, 82]. Secondary metabolism of *Streptomyces* is regulated by a wide variety of small molecular-weight substances and regulatory proteins [82, 115].

Several *S. griseus* metabolites have been identified and characterised, in particular the well-known antibiotic streptomycin [111, 112, 116]. Secondary metabolites have a wide range of biological activities. The most important is antibiotic activity i.e. growth inhibition or killing of other organisms. Antibiotic activity can include toxic effects against multicellular organisms such as invertebrates and plants as well as antibacterial and antifungal activities. Examples of such compounds include cycloheximide (**1**) [117-119] and candicidin (**2**) [120, 121] with antifungal properties, gephyromycin (**3**) [122] shows activity on neuronal cells and antibacterial compounds include frigocyclinone (**4**) [123] and feigrisolide C (**5**) [124]. Other biological activities of secondary metabolites include inhibition of spore germination, iron-uptake, selective transport of ions (as in the case of ionophore antibiotics), hormone-like roles in differentiation, immunosuppressive, herbicidal and participation in the regulation of primary metabolism. Secondary metabolites of *S. griseus* have also been shown to have insecticidal activity, valinomycin (**6**) [125] antioxidants such as diphenazithionin (**7**) [126] and anti-tumour antibiotics including fredericamycin A (**8**) [127], chromomycin A_3 (**9**) [128] and nonactin (**10**) [129] , Streptomycetes produce a large number of secondary metabolites and antibiotics represent only 10% of the total [82, 114]. Chemical structures of some of the secondary metabolites are shown in the Figure 1-7.

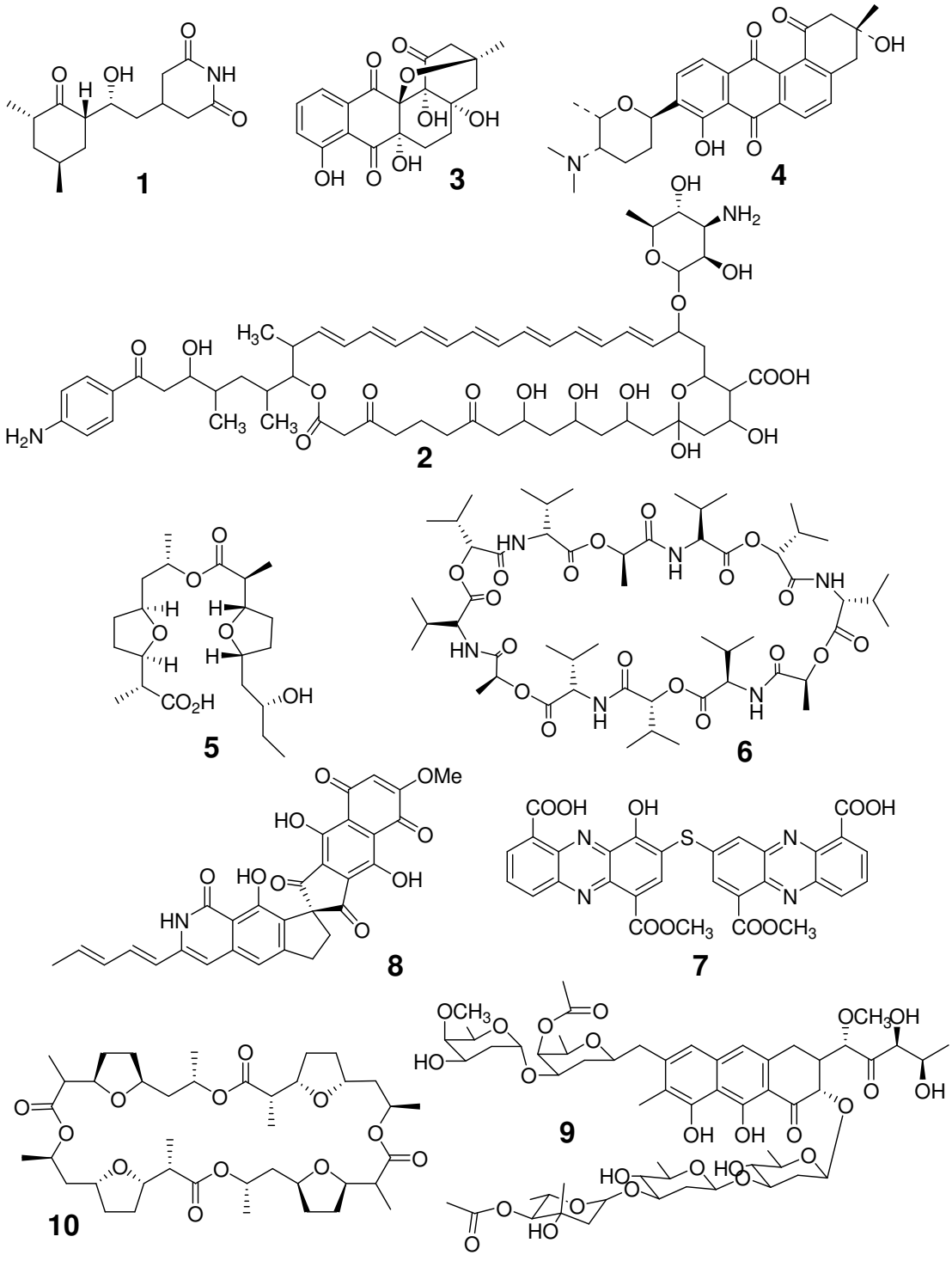
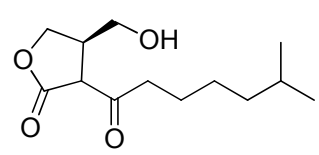


Figure 1-7 Structures of some of the important secondary metabolites of *S. griseus*

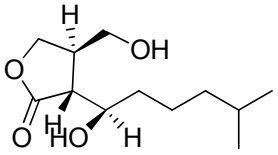
Various factors affect the production of secondary metabolites and antibiotics. Inorganic phosphate has an inhibitory effect on the biosynthesis of many antibiotics [130-132]. On the other hand, *S. griseus* growth has been found to be stimulated at high phosphate concentration [133]. The production of polyene antibiotics has been found to be particularly sensitive to phosphate in the production medium. One such example is the antibiotic candicidin [121, 134, 135]. A correlation between NAD(P) glycohydrolase activity and the production of antibiotics has also been reported [134, 136, 137]. Macrotetrolide antibiotics are produced by *S. griseus* under conditions that result in high NAD(P) glycohydrolase activity [138, 139]. A similar correlation was reported for the production of other antibiotics including viomycin and streptomycin both of which are produced by *S. griseus* [134]. Several of the secondary metabolites known to be produced by *S. griseus* are discussed in further detail in the following sections.

1.9.1 A-Factor

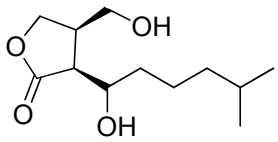
Metabolites with a γ-butyrolactone ring function as autoregulatory factors or microbial hormones for the expression of various phenotypes [140, 141]. A-factor is representative of γ-butyrolactone autoregulators [142]. Figure 1-8 shows the structure of A-factor, along with the structures of several other γ-butyrolactone autoregulators.

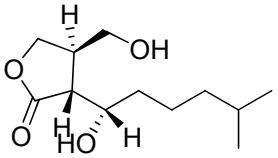


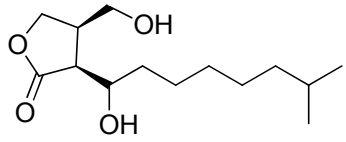
A-Factor from *S. griseus*

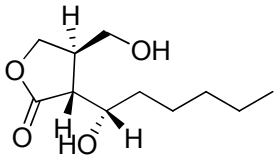


Factor 1 from *S viridochromogenes*

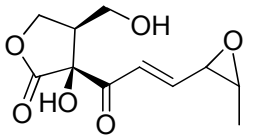




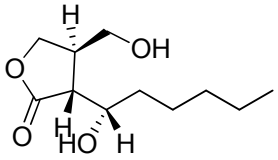




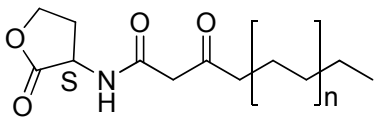
A-Factor from *S. bikiniensis* and S. cyaneofuscantus



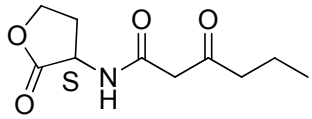
Butalactin from *Streptomyces* sp. Y-86,36923



Virginiae butanolides from *S. virginiae*



Conjugation factor of Agrobacterium



Auto-inducer from *V.fischeri*

Figure 1-8 Structure of autoregulatory factors having a γ-butyrolactone ring from various bacteria [143]

A-factor is essential for streptomycin production [144-146]. A-factor was originally reported by Khokhlov *et al*. in 1967. It is not only required for antibiotic production, but also morphological differentiation in *S. griseus* [147-149]. *S. griseus* produces a diffusible yellow pigment containing a sugar moiety when cultured on phosphate-depleted agar. Production of this yellow pigment is also under the control of A-factor (an A-factor-deficient strain does not produce the yellow pigment). The yellow pigment was found to be produced only in media containing low phosphate concentration. Thus, the production of the yellow pigment is switched-on by A-factor and controlled by phosphate levels in the medium [145]. A-factor exerts its regulatory

function at an extremely low concentration (as low as 10^{-9}M) [150]. Exogeneous supplementation or over-production of A-factor through genetic manipulation does not enhance streptomycin production [143, 151, 152]. However, if A-factor is introduced at the time of inoculation, streptomycin production starts a day early [143, 150, 153, 154]. The structure-activity relationship for A-factor has been investigated [145]. Virtually all alterations to its structure leads to a decrease or total disappearance of regulatory activity [143, 155]. In conclusion, A-factor is best thought of as a bacterial hormone, comparable to eukaryotic hormones such as the sex pheromones controlling differentiation in fungi [145].

1.9.2 Candicidin

Candicidin was first described as a metabolite of *S. griseus* by Lechevalier *et al*. in 1953 [156]. It was originally named C135, although it was renamed candicidin because of its strong activity against species of *Candida*. It is produced by *Streptomyces griseus* IMRU 3570 [120, 121].

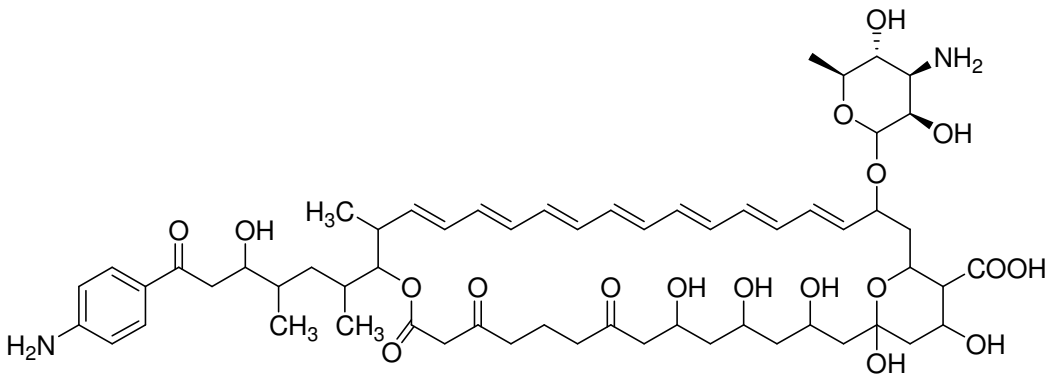


Figure 1-9 Structure of Candicidin-D

Candicidin is a heptaene antibiotic that contains the amino sugar mycosamine and the aromatic moiety 4-aminoacetophenone (Figure 1-9) [121].

1.9.3 Streptomycin

Streptomycin is the most famous of the antibiotics produced by *S. griseus*. Streptomycin has three constituents. A central 6-deoxyhexose component is

connected to N-methyl-L-glucosamine and scyllo-inositol-derived aminocyclitol components (Figure 1-10) [157, 158]. A cluster of about 30 coregulated genes in *S. griseus* directs streptomycin biosynthesis [159].

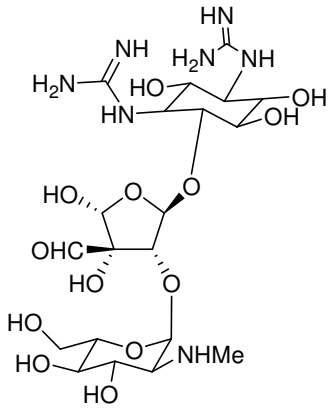
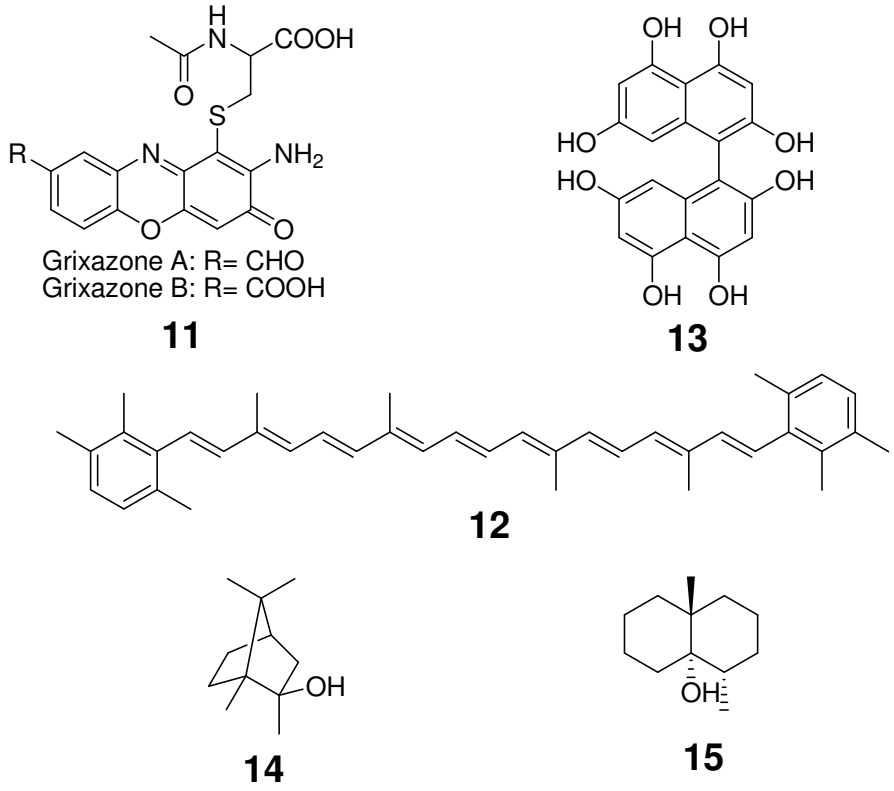


Figure 1-10 The structure of Streptomycin

Streptomycin was first isolated in 1943 by Albert Schalz, a graduate student at Rutgers University working under the supervision of Selman Waksman [112, 160]. It was the first antibiotic used to cure tuberculosis. Streptomycin stops bacterial growth by damaging cell membranes and inhibiting protein synthesis. Specifically it binds to the 23S subunit of the bacterial ribosome. This prevents the release of the growing polypeptide chain. Streptomycin is not toxic to humans because human ribosomes are structurally different to bacterial ribosomes.

Analysis of the genome sequence of *S. griseus* IFO 13350 strain [79] has revealed total of 34 gene cluster or genes for putative secondary metabolites. The products of six of these gene clusters is previously known which include streptomycin, grixazone (**11**) [161], a carotenoid (isorenieratene,**12**) [162, 163], the spore pigment HPQ (hexahydroxyperylenequinone) melanin (**13**) [164], alkylresorcinol and 2-methylisoborneol (**14**) [165]. There are a further 28 gene clusters or genes for putative secondary metabolites [79]. *S. griseus* genome contains many putative PKS, nonribosomal peptide synthetases (NRPS) and PKS-NRPS hybrid genes. Based on the sequence homology with other *Streptomyces* known genome sequences some gene cluster or gene products have been predicted including the earthy odorant geosmin (**15**) [39, 166], some siderophores, terpene compounds, hopanoid and lantibiotics [79].



Cyanobacteria and Actinomycetes are known to produce a number of secondary metabolites, including compounds that can impart an off-flavour to water and fish. Such compounds include methylisoborneol (MIB) and geosmin (Figure 1-11) [57, 145].

Figure 1-11 Structures of geosmin (A) and methylisoborneol (MIB) (B)

Off-flavours associated with these compounds have been reported in Japan, the USA, Europe and many other developed countries [82, 114, 134]. The occurrence of MIB, geosmin and microcystins has been assessed in 485 catfish production ponds in USA [115]. It was found that over 47% of ponds had detectable concentrations of these compounds. Microcystin were found in less than 10% of ponds at a concentration higher than 1ng/ml [114]. MIB and/or geosmin occurred in 25% of the samples [22]. Exposure of rainbow trout hepatocytes to geosmin and MIB for 48h at 15^{o}C resulted

in a reduction in cell viability in the presence of MIB, but not geosmin. DNA damage was observed at 0.45 mg/L/10 mg/L of geosmin/MIB, respectively. It has been suggested that these compounds can produce DNA breaks in fish cells which could lead to harmful mutations e.g. cancer [115].

Aims and objectives

The aim of this project was to attempt to identify the ichthyotoxic metabolites produced by the actinomycetes belonging to the S. *griseus* clade associated with the fish mortality incidents. The principal objectives were:

1) To culture the isolate producing the highest levels of ichthyotoxic metabolites (actino 13) and investigate whether ichthyotoxic metabolites can be extracted using organic solvents from the biomass or the culture supernatant.
2) To use HPLC to fractionate extracts containing ichthyotoxic metabolites and test the bioactivity associated with each HPLC fraction.
3) To use LC-MS to investigate the composition of HPLC fractions that exhibit ichthyotoxic activity.
4) To attempt to identify the purified ichthyotoxic metabolites using mass spectrometry and NMR spectroscopy.
5) Investigate the production of ichthyotoxic metabolites by the *Streptomyces griseus* DSM 40236.

.

Part I: Investigation of fish-toxins produced by *Streptomyces* species

CHAPTER 2 - RESULTS & DISCUSSION

2.1 Fish-toxin-producing strains

Six different strains were provided to us by Dr J. Parry of Lancaster University as shown in Table 2-1. These strains were collected from sediments of waters where large-scale fish mortalities had taken place. All of these strains were reported to belong to the *Streptomyces griseus* clade [19].

Table 2.1 Fish toxin-producing strains belonging to the *S. griseus* clade

	Strain		Strain
1	Actino 2	4	Actino 17
2	Actino 12	5	Strain 62
3	Actino 13	6	Strain 82

2.2 LC-MS profiling of metabolites produced by each strain

Spore suspensions were prepared from confluent lawns of each strain grown on SFM agar plates. SMM liquid medium was separately inoculated with spores of each of the six strains and the resulting cultures were incubated for 6 days at 30°C. Culture supernatants were separated from the biomass by centrifugation and extracted with ethyl acetate. The extracts were concentrated using a rotary evaporator and dissolved in 80% methanol for LC-MS analysis.

The base-peak chromatograms obtained from LC-MS analysis of the organic extracts of the culture supernatants of each of the six strains are shown in Figure 2-1.

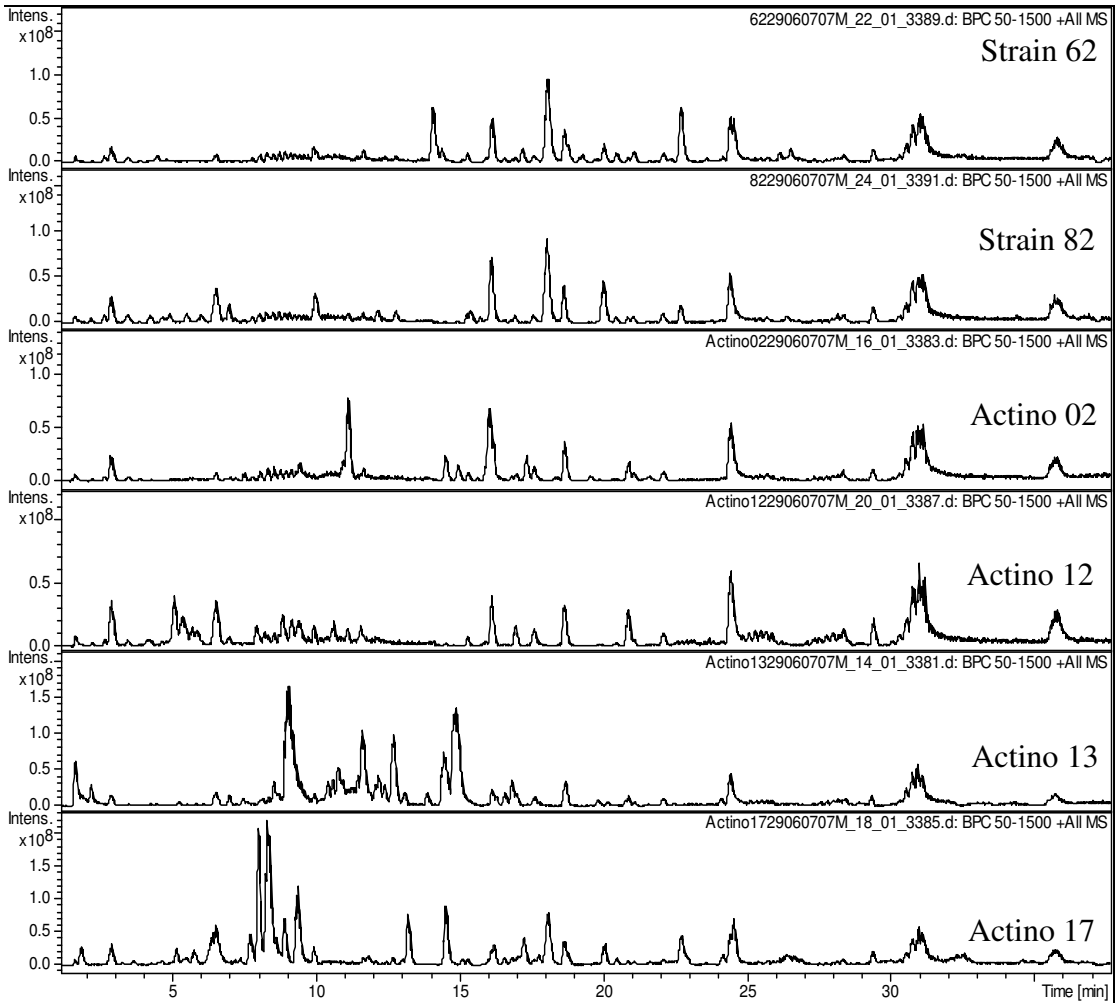


Figure 2-1 Base peak chromatograms from LC-MS analyses of organic extract of the agar media of the six toxin-producing Actinomycetes strains

It is reasonable to assume that the ichthyotoxic metabolites are common to all six strains. On comparison of the LC-MS data there were several *m/z* values found to be common in the organic extract of the agar media of the six strains. A few of these are shown in the table below;

Table 2.2 List of a few of the compounds common to all the organic extracts of the agar media from cultures of the six actinomycetes

Retention time (min)	Major ion (*m/z*)	Retention time (min)	Major ion (*m/z*)
2.72	295	30.42	473
6.36	278	30.61	429
18.55	274	30.82	385
24.25	258	31.02	341
29.18	358	35.58	369,413

m/z values of these species were searched in the scientific literature using chemical abstract database (SciFinder) to see any relevance to a known metabolite of *Streptomyces griseus*. As a result, the compound with *m/z* 258 ($M + H^+$) was found to correspond to the known antibiotic Indolmycin (Figure 2-2) [168, 169] with a reported molecular weight of 257.3. No other matches of observed *m/z* values to known compounds were identified.

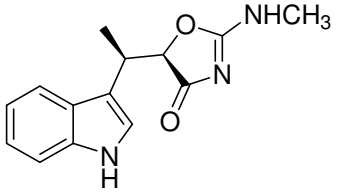


Figure 2-2 Chemical structure of Indolmycin presumably being produced by the six Actinomycetes strains

Parry *et al*., have reported previously that the biological activity in culture supernatants of the Actino 13 strain against fish was greater than the activity of the culture supernatants from the other strains. Thus we decided to concentrate solely on the Actino 13 strain for all subsequent experiments.

2.3 16S rRNA sequence analysis of the Actino 13 strain

Actino 13 was classified as belonging to the *Streptomyces griseus* clade during early research published by the Environment Agency [19, 170]. Thus we purchased a type strain of *Streptomyces griseus* (DSM 40236) to compare with the Actino 13 strain. The DSM 40236 strain was phenotypically different from the Actino 13 strain when grown on SFM agar media. The mycelium as well as the spores of the Actino 13

strain were always darker in colour than the mycelium and spores of the *S. griseus* DSM 40236 strain. Thus it was decided to compare the 16S rRNA sequences of the Actino 13 and DSM 40236 strains.

Genomic DNA from the *Streptomyces griseus* DSM 40236 and Actino 13 strains was isolated using a standard protocol [30]. PCR primers designed specifically to amplify the 16S rRNA gene of *Streptomyces griseus* were purchased from a commercial supplier and used in PCR reactions with genomic DNA from the DSM 40236 and Actino 13 strains as a template [171]. The products of these reactions were analysed by agarose gel electrophoresis. There was no product band of the expected size (1.5 kb) for the Actino 13 strain but there was a product of molecular size ~1.5 kb from the DSM 40236 strain (Figure 2-3).

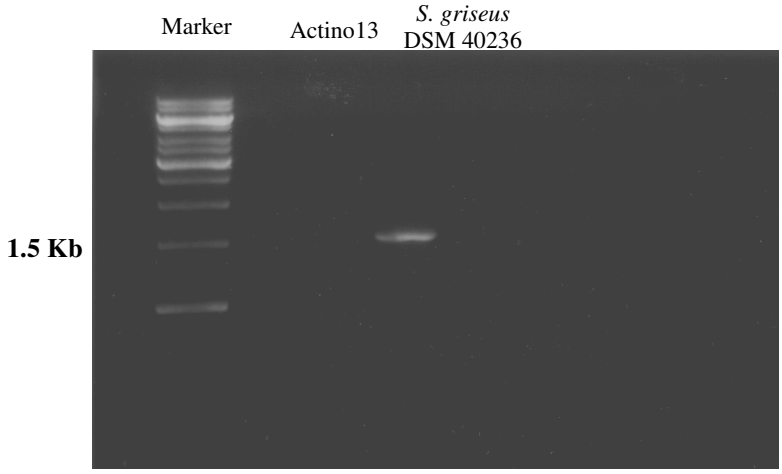


Figure 2-3 1% Agarose gel electrophoresis analysis of PCR products obtained using *S. griseus* 16S rRNA specific primers and genomic DNA from the DSM 40236 and Actino 13 strains as templates

To overcome this problem, a second set of PCR primers designed to amplify 16S rRNA genes from diverse Actinomycete species [172] were used. The PCRs were analysed by agarose gel electrophoresis and products of the expected size (1.5 kb) for both strains were observed (Figure 2-4).

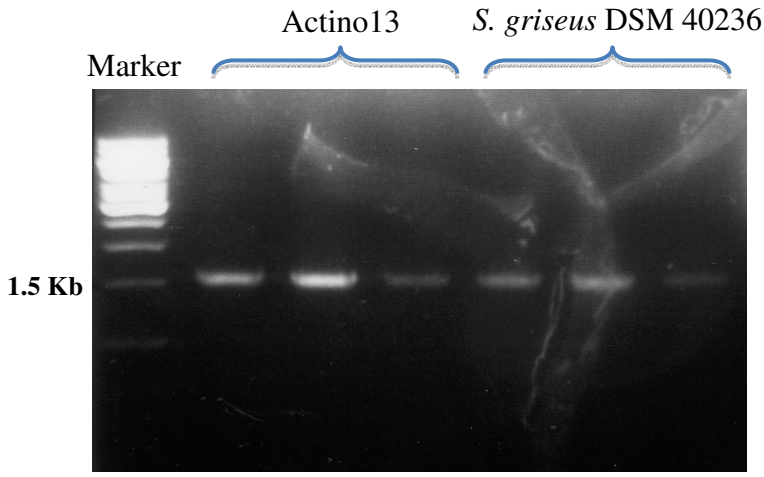


Figure 2-4 1% Agarose gel electrophoresis analysis of PCR products obtained using primers for amplification of 16S rRNA genes from diverse actinomycetes and genomic DNA from the DSM 40236 and Actino 13 strains as templates

The PCR products were purified from the gel and sequenced. Comparison of the sequences obtained indicated that the 16S rRNA genes of *S. griseus* DSM 40236 and Actino 13 strains are virtually identical (Figure 2-5). Thus the Actino 13 strain does appear to fall within the *S. griseus* clade as previously suggested by Parry and co-workers.

```
Actino13-seq      1 catgcaagtcgaacgatgaagcctttcggggtggattagtggcgaacgggtgagtaacacgtgggcaatctgcccttcactctgggacaagccctggaaacggggtctaa
DSM40236-seq      1 catgcaagtcgaacgatgaagcctttcggggtggattagtggcgaacgggtgagtaacacgtgggcaatctgcccttcactctgggacaagccctggaaacggggtctaa

Actino13-seq    111 taccggataacactctgtcccgcatgggacggggttaaaagctccggcggtgaaggatgagcccgcggcctatcagcttgttggtggggtaatggcctaccaaggcgacg
DSM40236-seq    111 taccggataacactctgtcccgcatgggacggggttaaaagctccggcggtgaaggatgagcccgcggcctatcagcttgttggtggggtaatggcctaccaaggcgacg

Actino13-seq    221 acgggtagccggcctgagagggcgaccggccacactgggactgagacacggcccagactcctacgggaggcagcagtggggaatattgcacaatgggcgaaagcctgatg
DSM40236-seq    221 acgggtagccggcctgagagggcgaccggccacactgggactgagacacggcccagactcctacgggaggcagcagtggggaatattgcacaatgggcgaaagcctgatg

Actino13-seq    331 cagcgacgccgcgtgagggatgacggccttcgggttgtaaacctctttcagcagggaagaagcgaaagtgacggtacctgcagaagaagcgccggctaactacgtgccag
DSM40236-seq    331 cagcgacgccgcgtgagggatgacggccttcgggttgtaaacctctttcagcagggaagaagcgaagtgacggtacctgcagaagaagcgccggctaactacgtgccag

Actino13-seq    441 cagccgcggtaatacgtagggcgcaagcgttgtccggaattattgggcgtaaagagctcgtaggcggcttgtcacgtcggatgtgaaagcccggggcttaaccccgggtc
DSM40236-seq    441 cagccgcggtaatacgtagggcgcaagcgttgtccggaattattgggcgtaaagagctcgtaggcggcttgtcacgtcggatgtgaaagcccggggcttaaccccgggtc

Actino13-seq    551 tgcattcgatacgggctagctagagtgtggtaggggagatcggaattcctggtgtagcggtgaaatgcgcagatatcaggaggaacaccggtggcgaaggc-ggatctct
DSM40236-seq    551 tgcattcgatacgggctagctagagtgtggtaggggagatcggaattcctggtgtagcggtgaaatgcgcagatatcaggaggaacaccggtggcgaaaggcggatctct

Actino13-seq    660 gggccattactgacgctgaggagcgaaagcgtggggagcgaacaggattagataccctggtagtccacgccgtaaacgttgggaactaggtgttggcgacattccacgtc
DSM40236-seq    661 gggccattactgacgctgaggagcgaaagcgtggggagcgaacaggattagataccctggtagtccacgccgta-acgttgggaactaggtgttggcgacattccacgtc

Actino13-seq    770 gtcggtgccgcagctaacgcattaagttccccgcctggggagtacggccgcaaggctaaaactcaaaggaattgacgggggcccgcacaagcagcggagcatgtggctta
DSM40236-seq    770 gtcggtgccgcagctaacgcattaagttccccgcctggggagtacggccgcaaggctaaaactcaaaggaattgacgggggcccgcacaagcagcggagcatgtggctta

Actino13-seq    880 attcgacgcaacgcgaagaaccttaccaaggcttgacatataccggaaagcatcagagatggtgccccccttgtggtcggtatacaggtggtgcatggctgtcgtcagct
DSM40236-seq    880 attcgacgcaacgcgaagaaccttaccaaggcttgacatataccggaaagcatcagagatggtgccccccttgtggtcggtatacaggtggtgcatggctgtcgtcagct

Actino13-seq    990 cgtgtcgtgagatgttgggttaagtcccgcaacgagcgcaacccttgttctgtgttgccagcatgcccttcggggtgatggggactcacaggagactgccggggtcaact
DSM40236-seq    990 cgtgtcgtgagatgttgggttaagtcccgcaacgagcgcaacccttgttctgtgttgccagcatgcccttcggggtgatggggactcacaggagactgccggggtcaact

Actino13-seq   1100 cggaggaaggtggggacgacgtcaagtcatcatgccccttatgtcttgggctgcacacgtgctacaatggccggtacaatgagctgcgatgccgcgaggcggagcgaatc
DSM40236-seq   1100 cggaggaaggtggggacgacgtcaagtcatcatgccccttatgtcttgggctgcacacgtgctacaatggccggtacaatgagctgcgatgccgcgaggcggagcgaatc

Actino13-seq   1210 tcaaaaagccggtctcagttcggattggggtctgcaactcgaccccatgaagtcggagttgctagtaatcgcagatcagcattgctgcggtgaatacgttcccgggcctt
DSM40236-seq   1210 tcaaaaagccggtctcagttcggattggggtctgcaactcgaccccatgaagtcggagttgctagtaatcgcagatcagcattgctgcggtgaatacgttcccgggcctt

Actino13-seq   1320 gtacacaccgcccgtcacgtcacgaaagtcggtaacacccgaagccggtggcccaaccccttgtgggagggagctgtcgaaggtgggactggcgattgggac
DSM40236-seq   1320 gtacacaccgcccgtcacgtcacgaaagtcggtaacacccgaagccggtggcccaaccccttgtgggagggagctgtcgaaggtgggactggcgattgggac
```

Figure 2-5 Sequence alignment of Actino 13 and *S. griseus* DSM 40236 16S rRNA genes. The sequences are >98% identical

2.4 Analysis of major metabolites in extracts of Actino 13 culture supernatants

The organic extracts of Actino 13 culture supernatants were separated by HPLC monitoring absorbance at 210 nm. The most abundant compounds of the extracts were collected and analysed by high resolution mass spectrometry to determine their molecular formulae. The chromatogram obtained from the HPLC separation is shown in Figure 2-6 (*m/z* values for the major species associated with each peak are shown above the peaks).

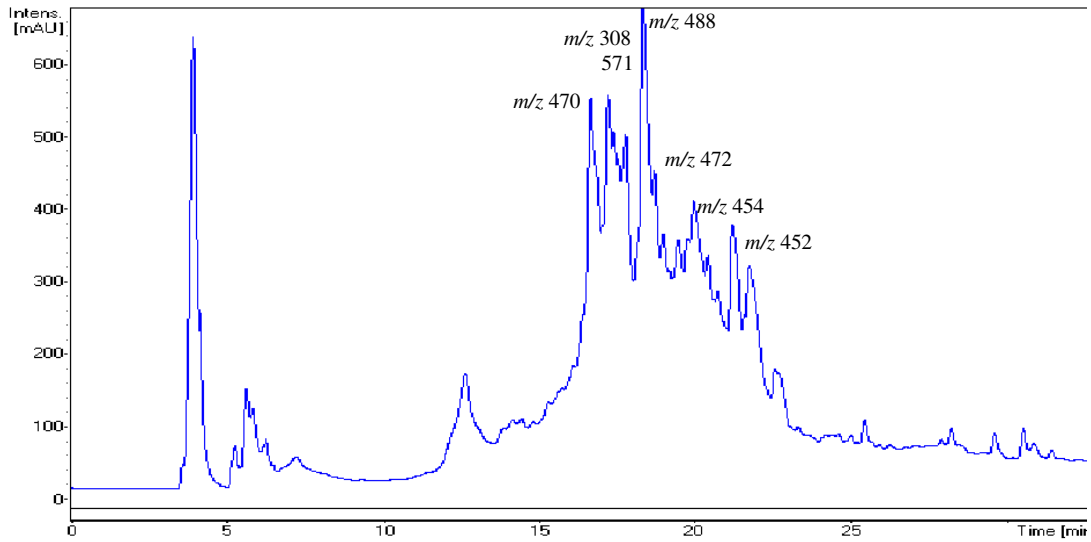


Figure 2-6 Chromatogram from HPLC separation of extracts of Actino 13 culture supernatants monitoring absorbance at 210nm

Table 2-3 lists the deduced molecular formulae of the collected compounds. These molecular formulae were used to search the Chemical Abstracts database (using SciFinder) to establish whether they correspond to the molecular formulae of known compounds. The results are summarised in Table 4. Results from the database search showed that a compound of *m/z* 308 has been reported very recently [173]. This natural compound, methyl phenatate A (Figure 2-7), has been reported to have identical molecular formula and is produced by a *Streptomyces* sp. H7372 isolated from mangrove soil of Sabah, Borneo, Malaysia. Another hit on the database search was of the compound with *m/z* 472. The match to this compound is a quinazolinone alkaloids cottoquinozoline C (Figure 2-7), which is one of the bioactive compound isolated and identified from coral-associated fungus *Aspergillus versicolor* LCJ-5-4 [174].

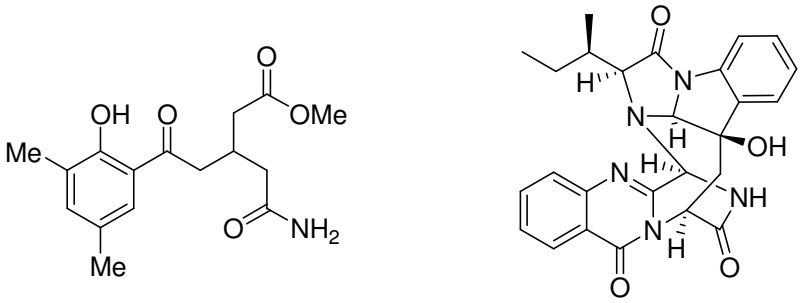


Figure 2-7 Chemical structures of methyl phenatate A and cottoquinozoline C

Compounds with high absorbance at 210 nm gave *m/z* value of 470, 308, 488, 472, 454, 452 and 571. Compound of *m/z* value of 472 might be the parent ion of *m/z* 454 because the difference between the two is 18Da which may correspond to loss of a water molecule. Similarly with the difference of 18Da, compound of *m/z* value 488 potentially be the parent ion of *m/z* 470 (Figure 2-8). MS/MS analyses of the collected compounds resulted, in most cases, in fragmentation to produce a common ion with *m/z* 144 (Figure 2-8). This suggests most of the compounds possess a common core structure that corresponds to the *m/z* = 144 ion. Fragmentation of the *m/z* = 144 ion resulted in a neutral loss of 44 Da, which could correspond to CO_2. Fragmentation of the resulting ion *m/z* = 100 ion resulted in a neutral loss of 28 DA, corresponding to CO. The *m/z* = 144 ion was analysed using TOF-MS to determine possible molecular formulae for it. One molecular formula that was consistent with the data obtained is $C_7H_{14}NO_2$. A search of the Chemical Abstracts database using this molecular formula returned more than 100 different compounds.

Further analysis of the reported compound, methyl phenatate A with similar *m/z* of 308, indicated that the fragmentation pattern may highly resemble to the compound of *m/z* 308 identified in the culture supernatant of Actino 13. Suspecetd fragmentation pattern is shown in the Figure 2-8 which highly resemble to the MS/MS data obtained for the compound with *m/z* 308 from Actino 13 (Figure 2-9).

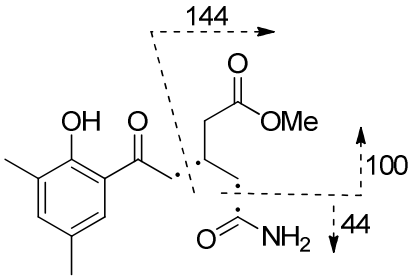


Figure 2-8 Suspected fragmentation pattern of methyl phenatate A

Table 2.3 Molecular formulae determined for some major metabolites in ethyl acetate extracts of the culture supernatants of Actino 13 strain and a summary of the results of searches of the Chemical Abstracts database for compounds with the deduced molecular formulae (using SciFinder)

Measured *m/z*	Deduced molecular formulae	Err (ppm)	SciFinder Search results
470.1828,	$C_{13}H_{32}N_3O_{15}$,	0.5	No hits
470.1809,	$C_{25}H_{28}NO_8$,	4.5	>400 hits
470.1823,	$C_{26}H_{24}N_5O_4$,	1.6	>500 hits (synthetic compounds including pesticides, anticancer, treatment of Parkinson's
470.1841	$C_{14}H_{28}N_7O_{11}$	2.3	disease. HIV, No hits
308.1492	$C_{16}H_{22}NO_5$,	3.35	1 hit (a natural compound, methyl phenatate A) [173]8 hits (synthetic
	$C_{17}H_{18}N_5O$	1.01	compounds) 6 hits (all synthetic compounds)
488.1928	$C_{26}H_{26}N_5O_5$	3.95	Over 900 hits
472.1993,	$C_{28}H_{28}N_2O_5$,	2.14	Over 1000 hits,
472.1979	$C_{26}H_{26}N_5O_4$	0.71	3 hits (including a natural compound Cottoquinazoline C [174] and a synthetic azaindoleoxoacetic piperazine derivative [175]
454.1887,	$C_{28}H_{26}N_2O_4$,	1.68	Over 1000 hits,
454.1874	$C_{26}H_{24}N_5O_3$	4.64	1 hit (a synthetic azaindoleoxacetic piperazine derivative)

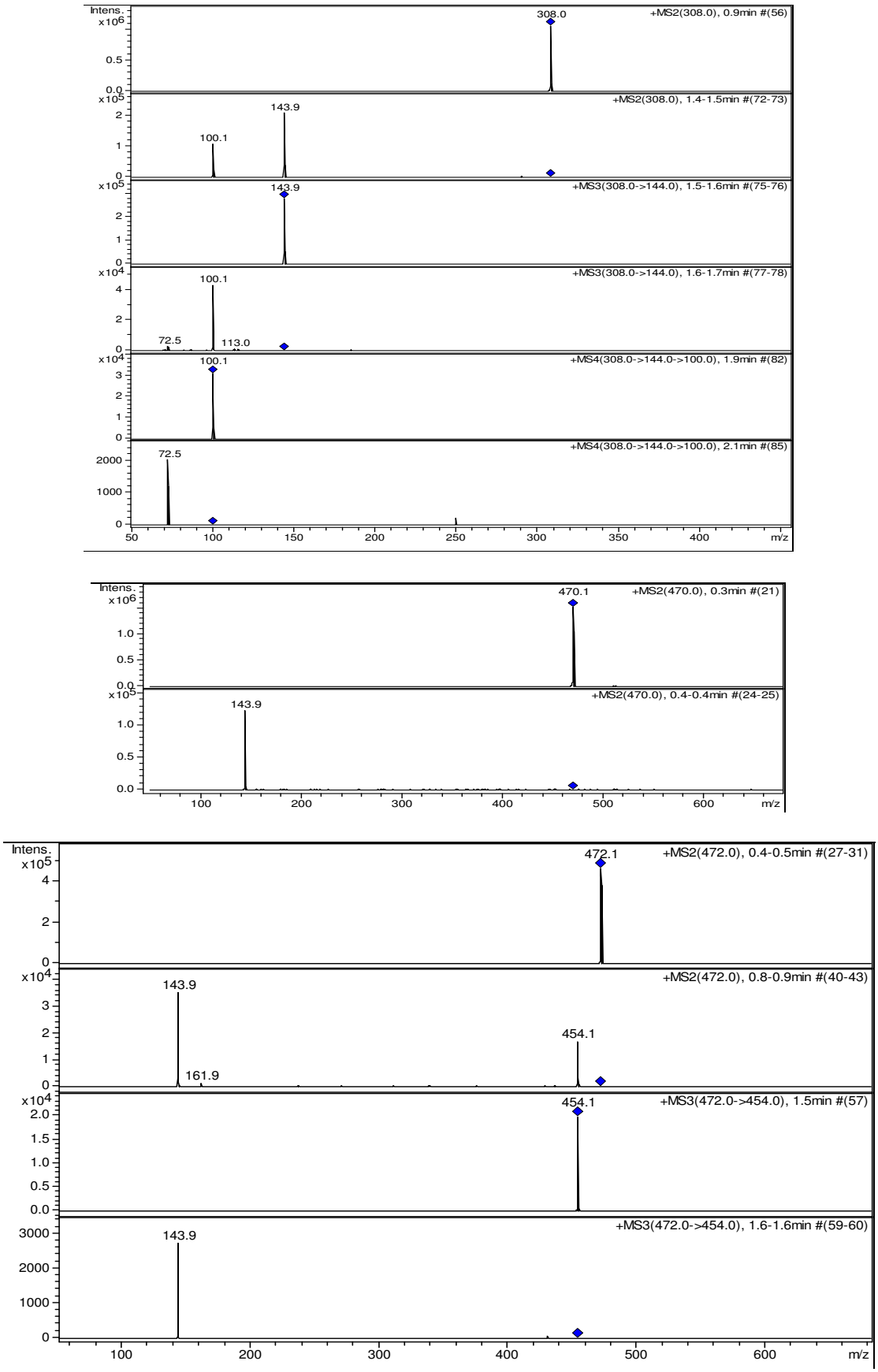


Figure 2-9 Fragmentation pattern of the compounds of mass 308, 470 and 472 Da to produce an ion of *m/z* 144

It is not possible to identify most of the major metabolites in Actino 13 culture supernatant extracts by HRMS because most of the molecular formulae correspond to hundreds or even thousands of known compounds (mostly synthetic). The two natural compounds which correspond to the two of the deduced molecular formulae have been reported very recently and there is no literature precedent of their ichthyotoxic activity.

2.5 Bioassay-guided HPLC fractionation of culture supernatant extracts

These experiments were carried out in collaboration with Professor John Lewis (Royal Holloway College London). All ichthyotoxicity tests of culture supernatant extracts and HPLC fractions were carried out by John Lewis and co-workers.

For the ichthyotoxicity tests, four samples were prepared from the Actino 13 and *S. griseus* DSM 40236 strains grown on starch casein agar medium. These were the organic filtrate, the aqueous filtrate, the organic extract and the aqueous extract (Figure 2-10).

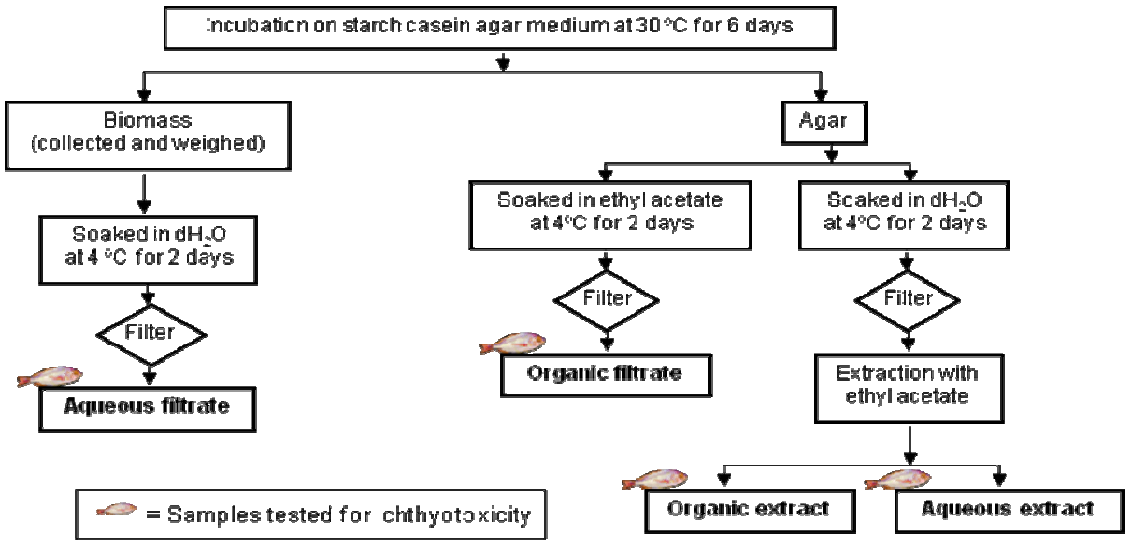


Figure 2-10 Scheme illustrating the sample preparation method from fish-toxin producing actinomycetes and the subsequent fractionations of the organic filtrate (1 of 4)

Samples were quantified on the basis of the amount of biomass produced in each culture. The organic filtrate, organic extract, aqueous filtrate and aqueous extract

samples at concentrations equivalent to 6, 12 and 18 mg of biomass were incubated separately with fish (carp fry or tench fry) for a total of 24 to 96 hours under controlled pH and temperature conditions.

Gills were removed from the fish, dehydrated, mounted on aluminium stubs, gold coated and examined by SEM [170]. The percentage of fusion of secondary lamellae and the quantity of microridging loss on the filamental epithelium of secondary lamellae were determined. The percentage of fusion of secondary lamellae for each sample is shown in Figure 2-11.

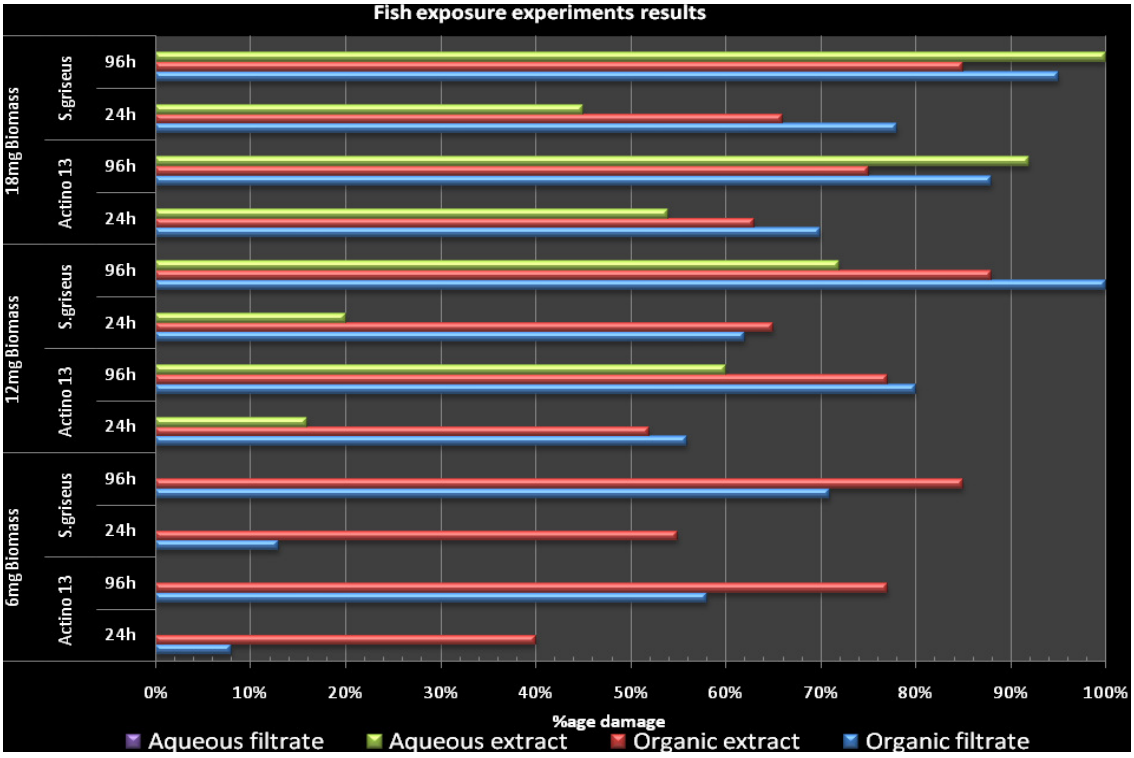


Figure 2-11 Graphical presentation of the results of the ichthyotoxicity tests. The x-axis corresponds to the percentage of fusion of secondary lamellae in each gill examined

After 24 hours of exposure, the organic extract sample corresponding to 6mg of biomass caused 40% damage to the gills whereas the corresponding organic filtrate sample caused 8% damage. Similar results were seen with the corresponding samples from the type strain *S. griseus* DSM40236. Gill damage after 24 hours exposure increased with the organic extract and organic filtrate samples corresponding to 12 mg of biomass. The aqueous extract also began to show some activity after 24 hours exposure with the sample equivalent to 12 mg of biomass. The organic filtrate sample

from the DSM40236 strain equivalent to 12 mg of biomass caused 100% damage to gills after 96 hours of exposure.

With the samples equivalent to 18mg of biomass, extensive gill damage resulted from exposure to the organic extract, organic filtrate and aqueous extract fractions from both strains. In the case of Actino 13 the damage was 70% with the organic filtrate, about 64% with the organic extract and about 54% with the aqueous extract after 24 hours of exposure. After 96 hours of exposure, 88% damage was caused by the organic filtrate, more than 75% damage occurred using the organic extract and more than 90% damage was observed using the aqueous extract. The pattern of damage was similar, but more severe, in the case of 96 hours exposure to samples from *S. griseus* DSM40236. Control experiments were also carried out along with the actual samples by using only the solvent, DMSO, as a negative control. No damage on fish gills was observed even after 96 hours of exposure.

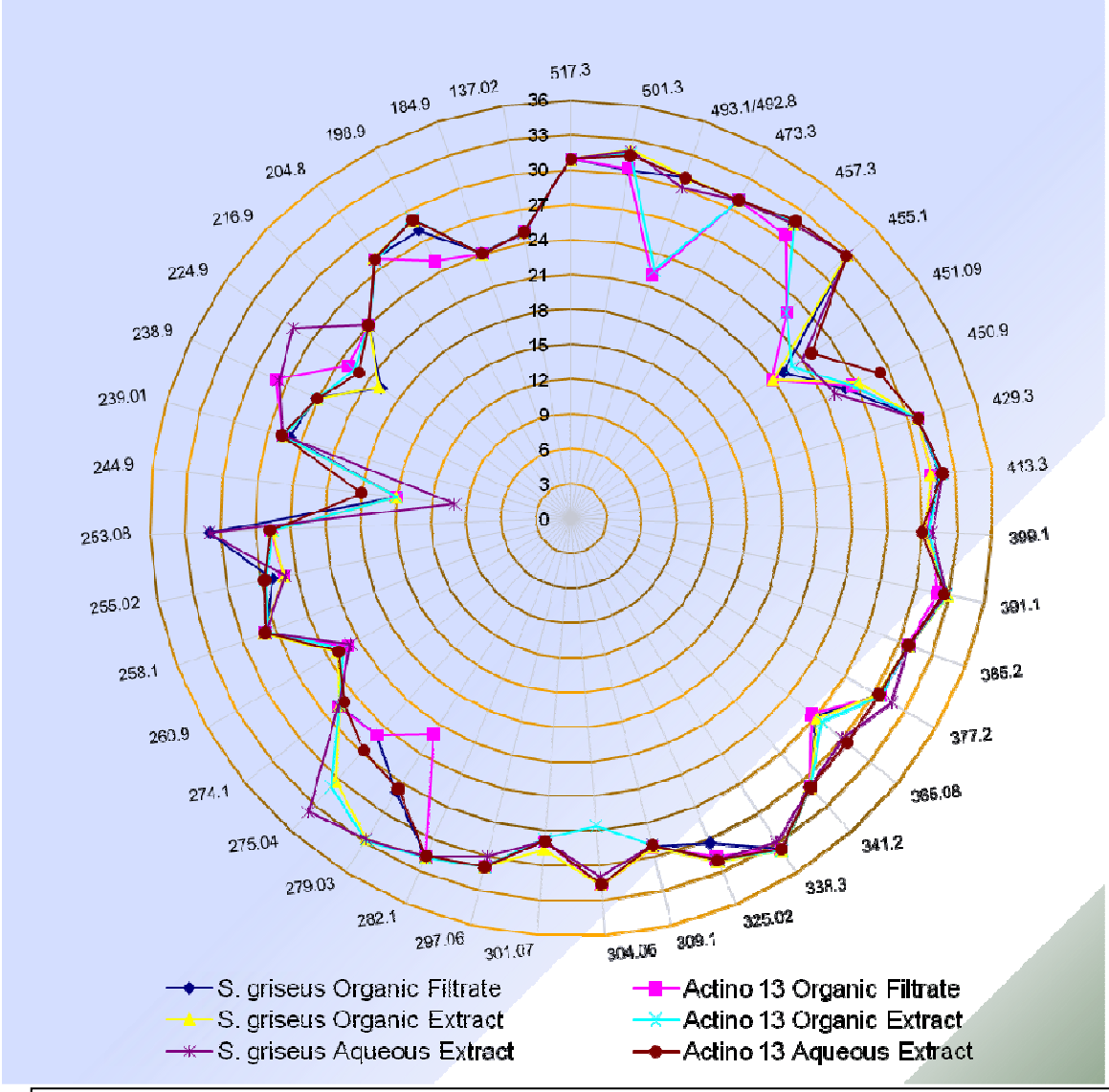
In summary, extracts from both Actino 13 and DSM40236 strains caused increasing gill damage with increasing concentration or exposure time. This indicates that the ichthyotoxic metabolites produced by both of the strains are likely to be the same. No gill damage was observed for the aqueous filtrate samples even after 96 hours of exposure at the highest concentration. Residual activity in the aqueous extract samples could be explained by incomplete extraction of the toxins into ethyl acetate during sample preparation. The finding that the *S. griseus* DSM40236 strain excretes metabolites that cause damage to fish gills is novel, and provides further support for the notion that the Actino 13 isolate belongs to the *Streptomyces griseus* clade.

Based on the results from the first phase of experiments we focused only on organic filtrate fraction of Actino 13 for the next phase of the project.

The above results showed that ichthyotoxic metabolites are present in the organic filtrate, the organic extract and the aqueous extract of both the *S. griseus* DSM40236 and Actino 13 strains. Thus, organic filtrate, organic extract and aqueous extract samples were prepared from both the Actino 13 and the *S. griseus* DSM 40236 strains, and analysed by LC-MS to attempt to identify metabolites that are present in

all of them. Retention times and *m/z* values for major components of each sample were compared (Figure 2-12).

Most of the compounds that are common to all samples eluted between 24 and 33 minutes. Many of the compounds with similar *m/z* eluted at slightly different times in all the samples. However few compounds, for instance *m/z* 184.9 and *m/z* = 137, eluted at approximately 24 and 25 minutes, respectively for all six samples (Figure 2-11). This comparative analysis was performed to have an idea of the extent of the number of common major compounds being produced by the *S. griseus* DSM40236 and the Actino 13 strains particularly in the most active samples (the organic filtrate and the organic extract, from the earlier ichthyotoxicity tests, Figure 2-11). The data in the chart also helped to design the sample preparation protocol for the next phase of ichthyotoxicity testing by indicating the retention times of majority of potential ichthyotoxic metabolites (common to all active sample) on a reverse phase C18 LC column. Analysis of the data in this way suggested that many compounds are present in all six samples and that identification of potential ichthyotoxic metabolites by such a comparative metabolic profiling method was likely to be challenging.



Numbers written by the concentric circles indicate retention times. Numbers around the outermost circle represent *m/z* values which were found to be common to the organic filtrate, organic extract and aqueous extract from both strains.

Figure 2-12 Comparison of the retention times for compounds giving common *m/z* values in all three extracts of both strains

To reduce the molecular complexity of the organic filtrate fraction from cultures of the Actino 13 strain, it was separated by HPLC (scheme shown in Figure 2-13) and fractions were collected every 3 minutes. The fractions were lyophilised and sent to Royal Holloway for biological testing. Five of the fractions were found to contain metabolites that cause damage to fish gills (Figure 2-14).

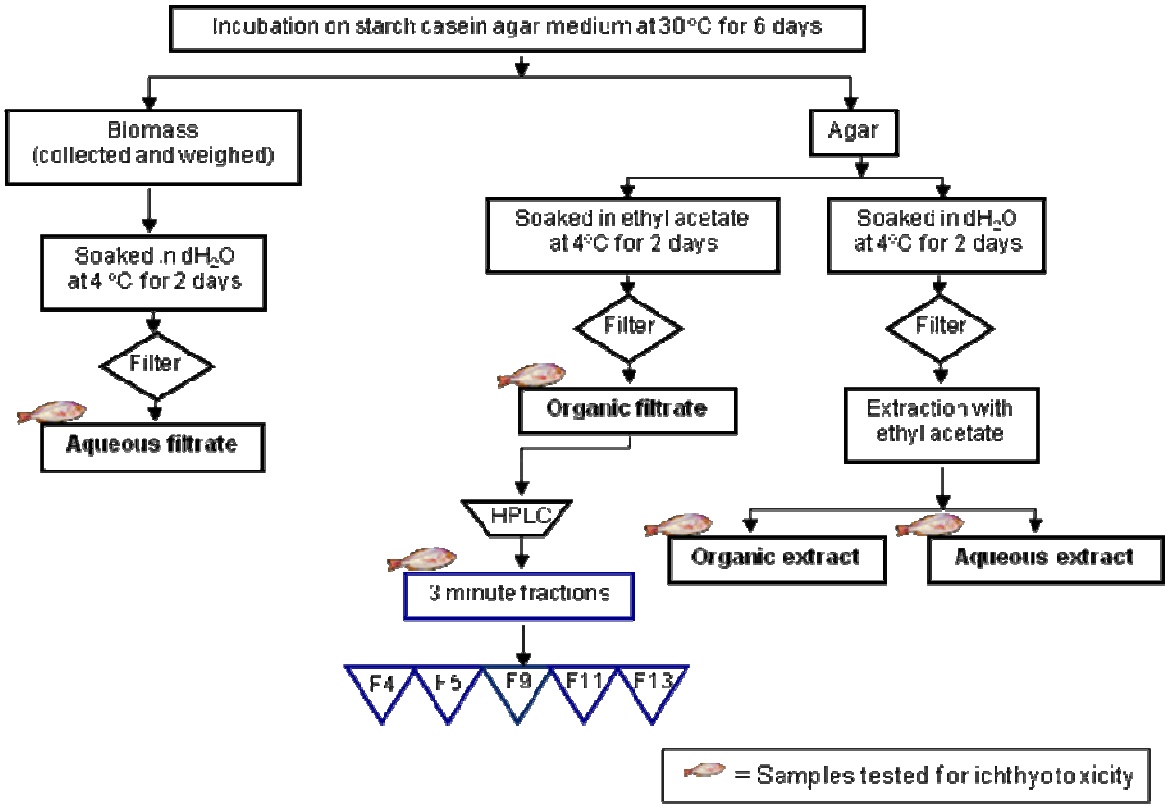


Figure 2-13 Scheme illustrating the sample preparation method from fish-toxin producing actinomycetes and the subsequent fractionations of the organic filtrate (2 of 4)

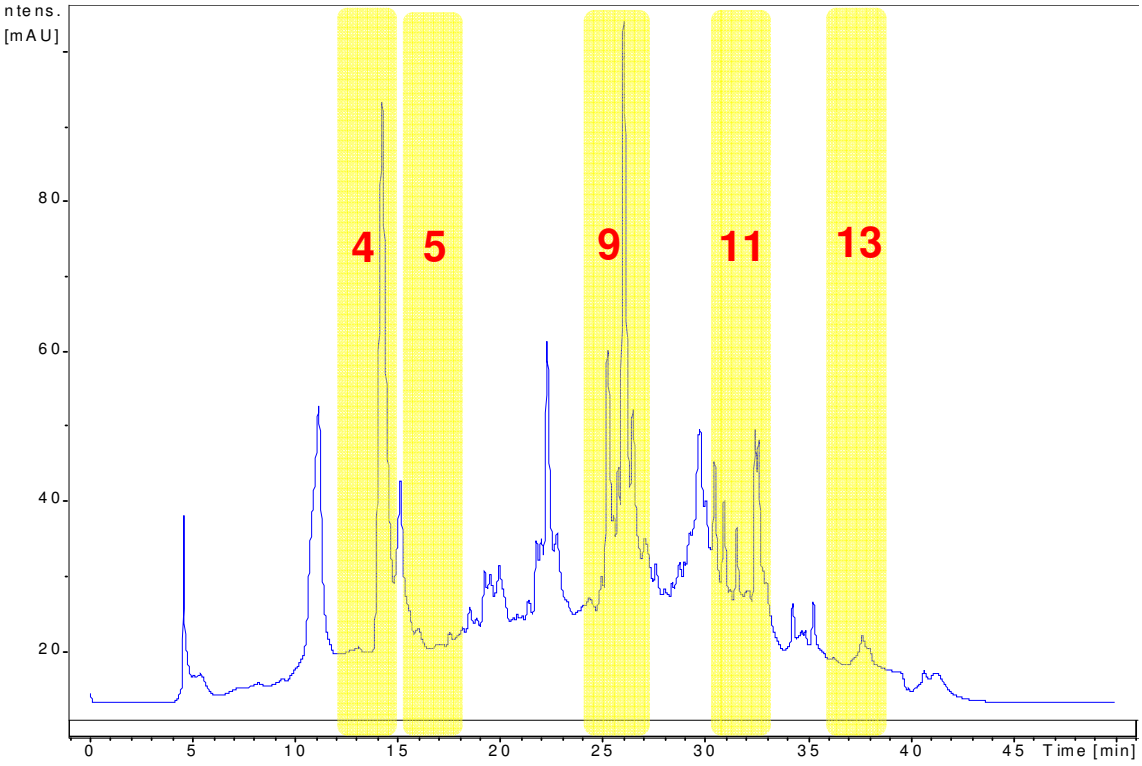


Figure 2-14 Chromatogram from HPLC separation monitoring absorbance at 275nm of the organic filtrate fraction from cultures of Actino 13. The time intervals corresponding to fractions 4,5,9, 11 and 13, all of which were found to contain ichthyotoxic metabolites, are highlighted

As well as exposing fish to single HPLC fractions, the fish were exposed to different combinations of fractions over 24-96 hours. The results of these experiments indicated that the fusion of secondary lamellae was significantly higher after 24 hours of exposure to a mixture of fractions 4/5/9 than after 24 hours exposure to the individual fractions (82.7% fusion with the mixture compared to 22.9-54.8% fusion for exposure to single fractions). Also a mixture of fractions 11 and 13 caused significantly more damage than each of these fractions in isolation (63.3% damage with the mixture compared to 40-43.8% damage from exposure to single fractions). After 96 hours of exposure, gill damage increased to 96.4% using a mix of fractions 4/5/9 compared to 77.8-82.9% damage with exposure to single fractions in contrast, a mix of fractions 11/13 did not cause significantly more damage than the isolated fractions after 96 hours of exposure (77.5% damage with the mix of fractions 11 and 13 compared to 66.7-75.9% damage from exposure to single fractions). Similar results were obtained when all 5 fractions (5, 6, 9, 11 & 13) were mixed but no significant additional synergy between the five fractions was observed (45.7% damage was observed after 24 hours and 90.6% damage was observed after at 96 hours exposure).

In summary, five different HPLC fractions were found to contain ichthyotoxic metabolites and mixtures of fractions had more biological activity than individual fractions. The mixture of fractions 4/5/9 showed highest activity and fraction 9 (F9) was the most active component of the mixture. These results clearly indicate that actino 13 does not produce a single ichthyotoxic metabolite, but rather it produces multiple metabolites that can cause damage to fish gills.

In the next phase of the project, the most active fraction from the initial HPLC separation, fraction 9 (F9), was further fractioned by HPLC and 5 minute fractions were collected and lyophilised (Figure 2-15). Each fraction was tested at Royal Holloway for its ability to cause damage to fish gills. Out of the eight fractions derived from F9 that were tested on fish, two (9F2 and 9F7) were found to be the most active (Figure 2-16).

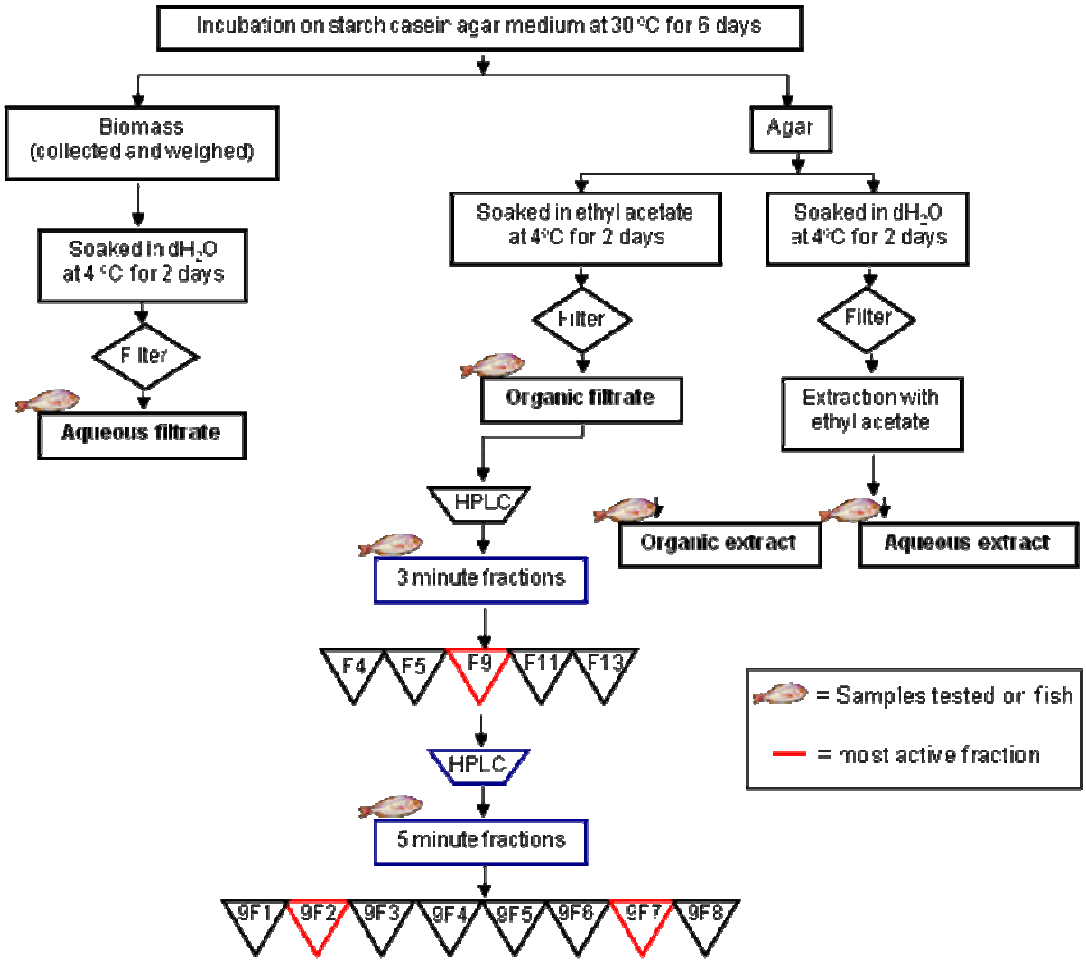


Figure 2-15 Scheme illustrating the sample preparation method from fish-toxin producing actinomycetes and the subsequent fractionations of the organic filtrate (3 of 4)

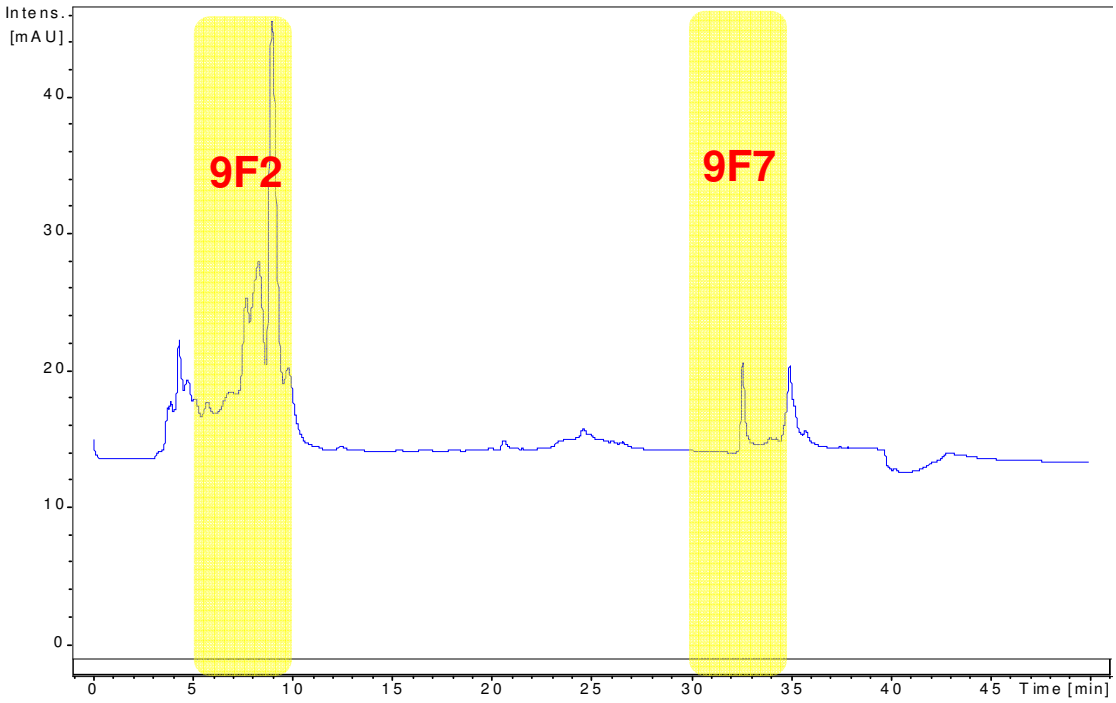


Figure 2-16 Chromatogram of fraction 9 from HPLC separation, monitoring absorbance at 275nm, from the initial separation of the Actino 13 organic filtrate

Fraction 9F2 was further fractioned by HPLC, monitoring absorbance at 267nm (Figure 2-17). Six fractions (9F2A – 9F2F) corresponding to the peaks indicated on the chromatogram were collected. These fractions were lyophilised and sent to Royal Holloway for fish-exposure experiments. Similarly fraction 9F7 was further fractioned by HPLC monitoring absorbance at 275nm (Figure 2-18). Fractions corresponding to peaks 9F7B, 9F7G, 9F7I, 9F7J and 9F7K were collected and, after lyophilisation, were sent to Royal Holloway for testing on fish. A schematic representing the sample preparation and all levels of HPLC fractionations is shown in Figure 2-19.

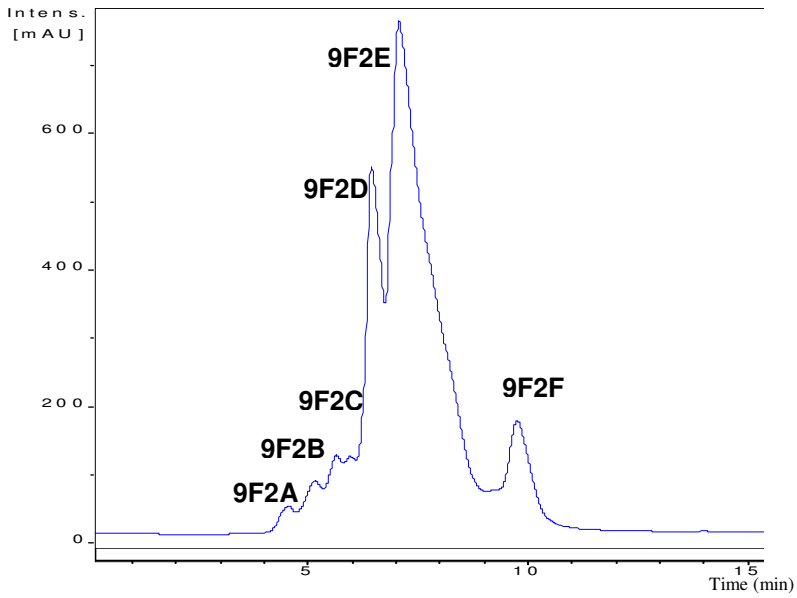


Figure 2-17 HPLC chromatogram of fraction 9F2 monitoring absorbance at 267nm

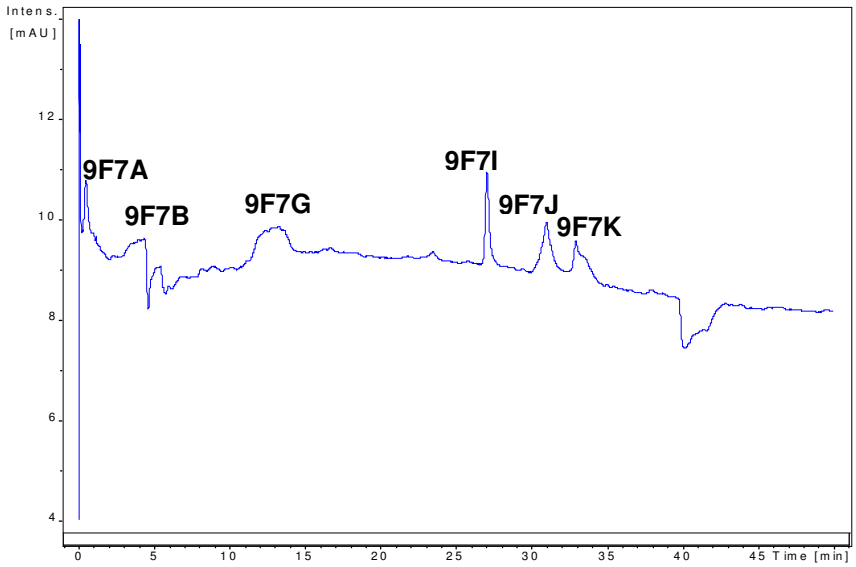


Figure 2-18 HPLC chromatogram of fraction 9F7 monitoring absorbance at 275nm

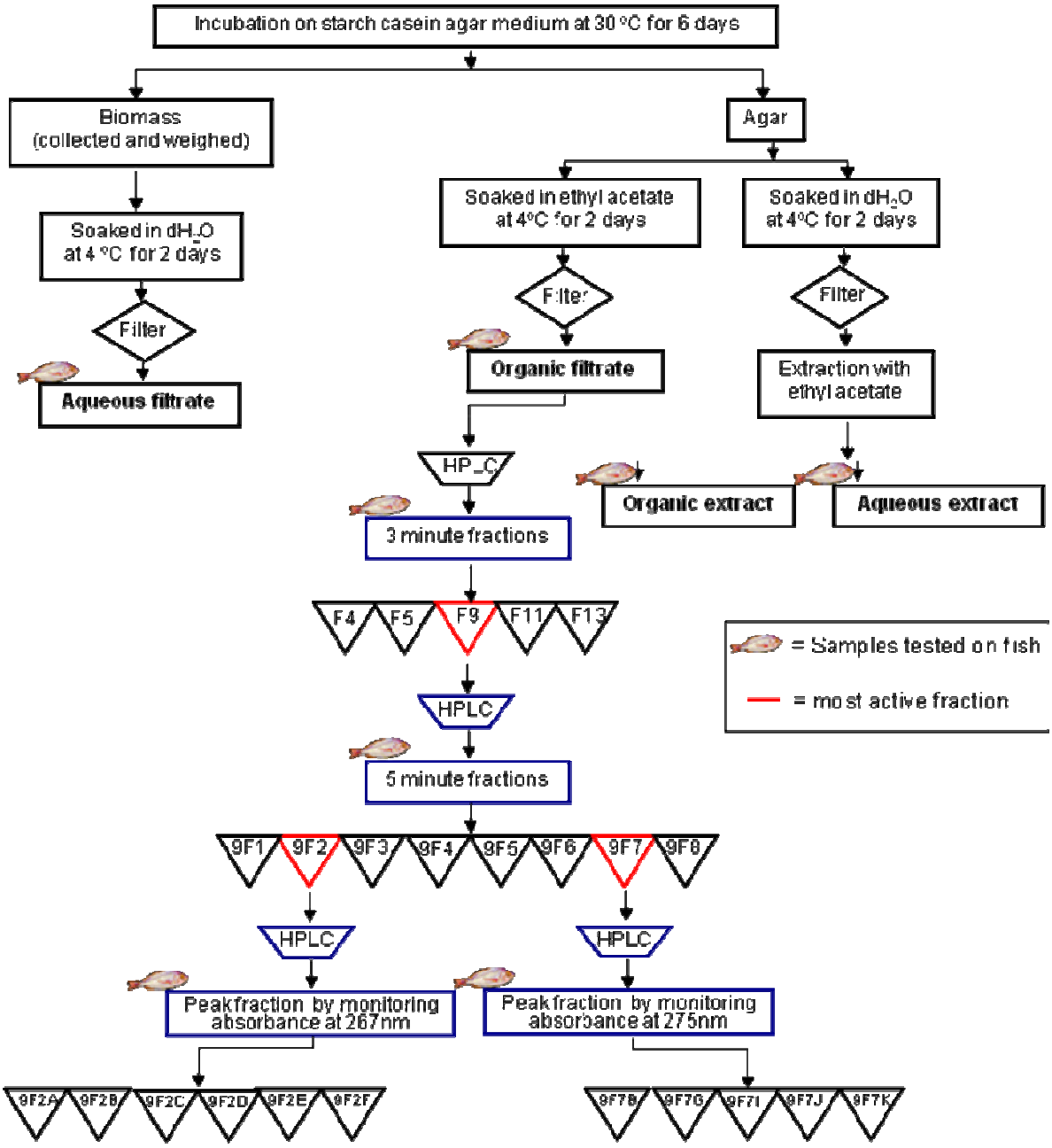


Figure 2-19 Scheme illustrating the sample preparation method from fish-toxin producing actinomycetes and the subsequent fractionations of the organic filtrate (4 of 4)

The results of the fish exposure experiments with fractions 9F2B, 9F2D, 9F2E, 9F2F, 9F7G, 9F7I, 9F7J and 9F7K are shown in Table 2-4. Fraction 9F2E had the highest activity. After only 24 hours of exposure, 74.2% of fusion of secondary lamellae was observed, with little loss of microridging on the filamental epithelium of secondary lamellae. Fusion of secondary lamellae increased to 90.3% after 96h exposure to fraction 9F2E and severe loss of microridging was observed. 9F2F was the second most active fraction. It caused 62.5% fusion of secondary lamellae and moderate loss of microridging on the filamental epithelium of secondary lamellae after 24h. 87.5%

fusion of secondary lamellae with still moderate loss of microridging was observed after 96 hours of exposure to this fraction. After 24 hours of exposure to fraction 9F7B 65.4% fusion of secondary lamellae with severe loss of microridging was observed. 84.6% fusion with severe loss of microridging occurred after 96 hours of exposure to this fraction. The other fractions tested also affected the secondary lamellae and filamental epithelium but to a lesser extent (Table 2-4).

Table 2.4 Results from exposure of fish to the major components of the 9F2 and 9F7 fractions

Fraction	24 hr exposure	96 hr exposure
9F2E	74.2%, +	90.3%, +++
9F2F	62.5%, ++	87.5%, ++
9F7B	65.4%, +++	84.6%, +++
9F2D	40%, ++	73.3%, ++
9F7I	23.3%, ++	43.3%, ++
9F7K	20.6%, ++	35.3%, ++
9F7J	18.9%, +	35.1%, ++
9F7G	13.8%, ++	31%, ++

% = fusion of secondary lamellae
\+ = little loss of microridging on the filamental epithelium of secondary lamellae
++ = moderate loss of microridging on the filamental epithelium of secondary lamellae
+++ = severe loss of microridging on the filamental epithelium of secondary lamellae

The compositions of the four most-active fractions (9F2E, 9F2F, 9F7B and 9F2D) were investigated using LC-MS (Figure 2-20). Each fraction was found to contain several compounds. The data were analysed in detail to establish the major molecular ions associated with each peak in each chromatogram (Figure 2-21–2-24).

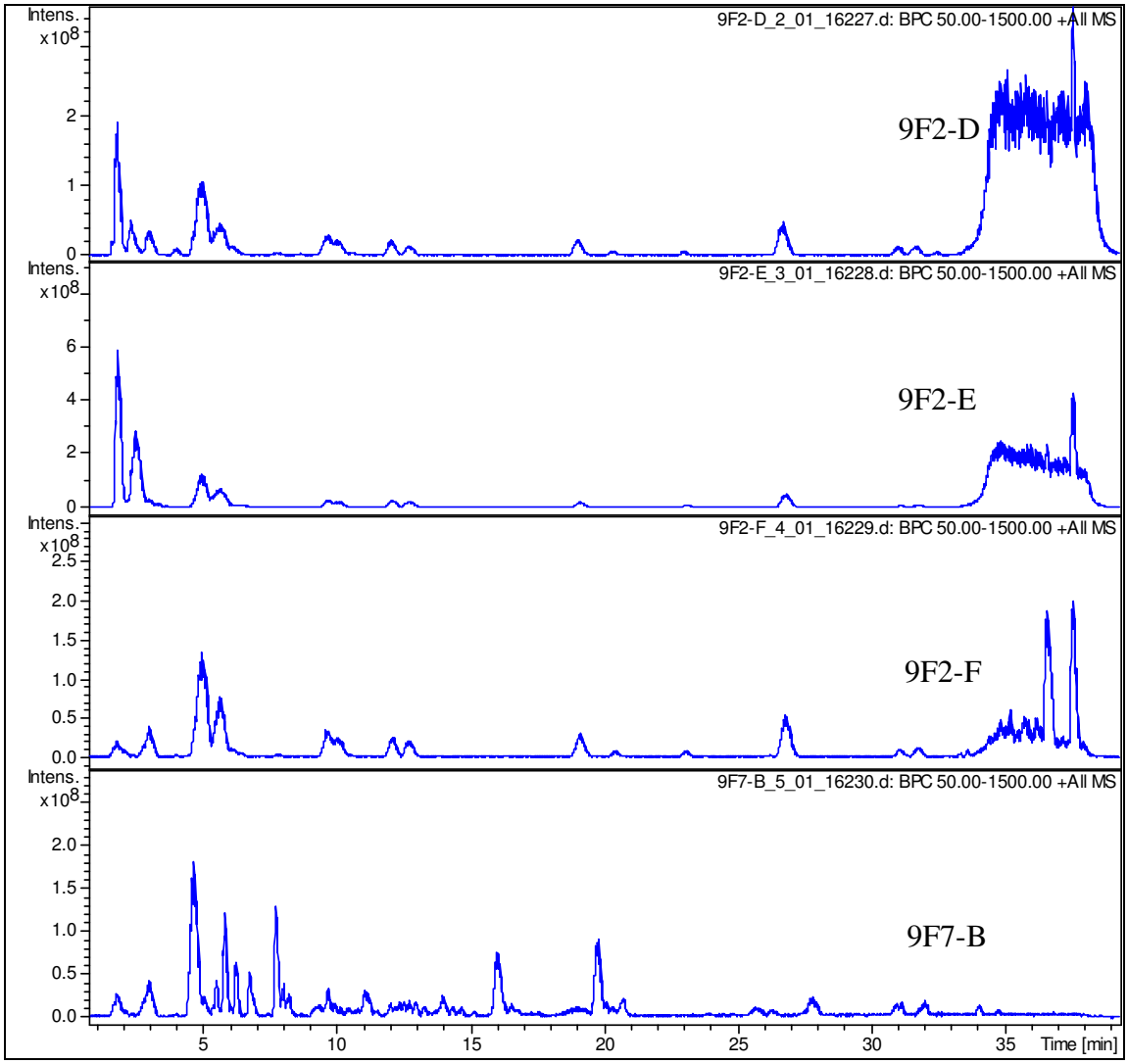


Figure 2-20 Comparison of base peak chromatograms from LC-MS analyses of the four most-active fractions 9F2E, 9F2F, 9F7B and 9F2D

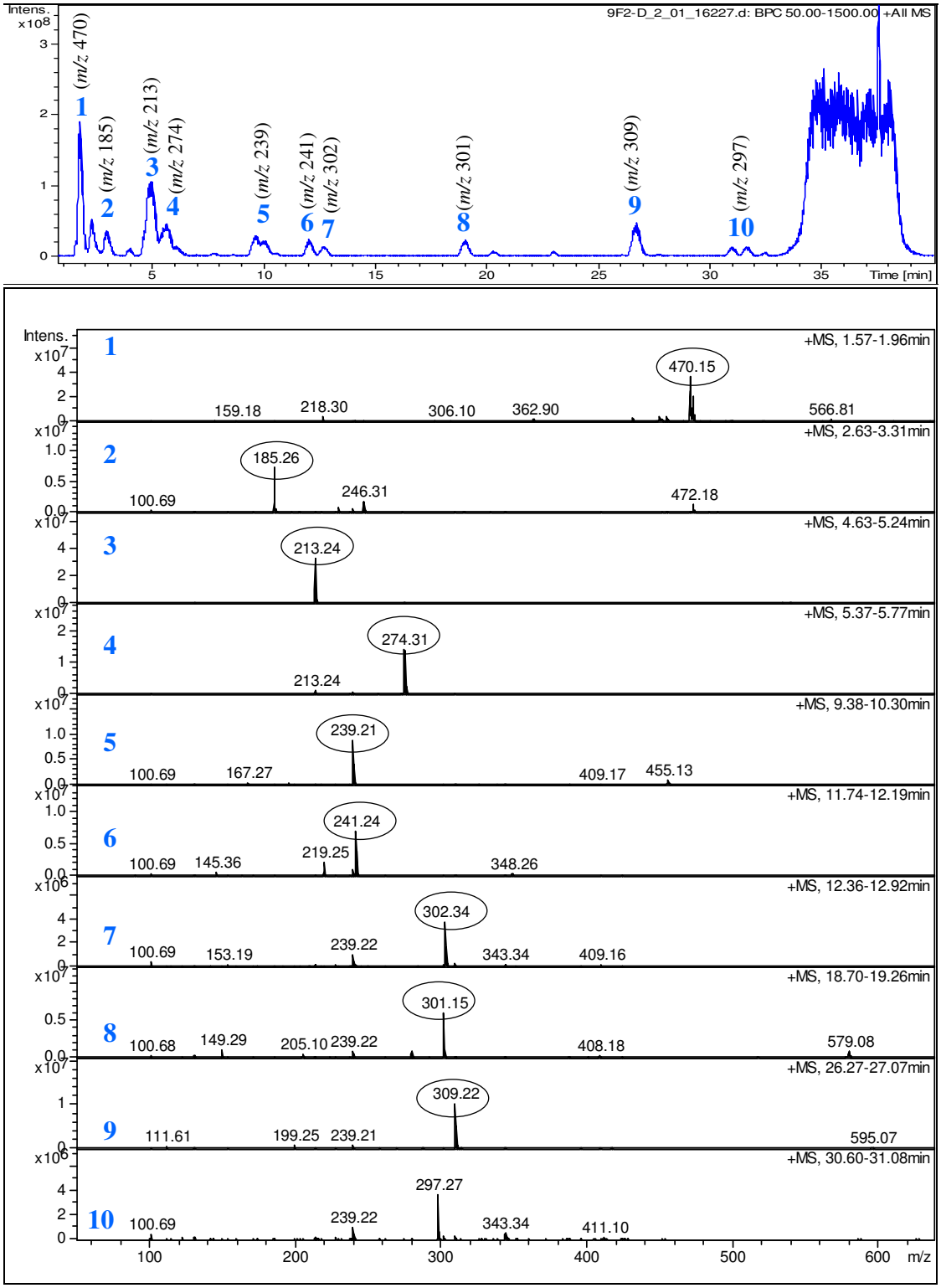


Figure 2-21 Chromatogram from LC-MS analysis of fraction 9F2D. Mass spectra for each of the numbered peaks are shown below the chromatogram and the major ion(s) in each spectrum are highlighted

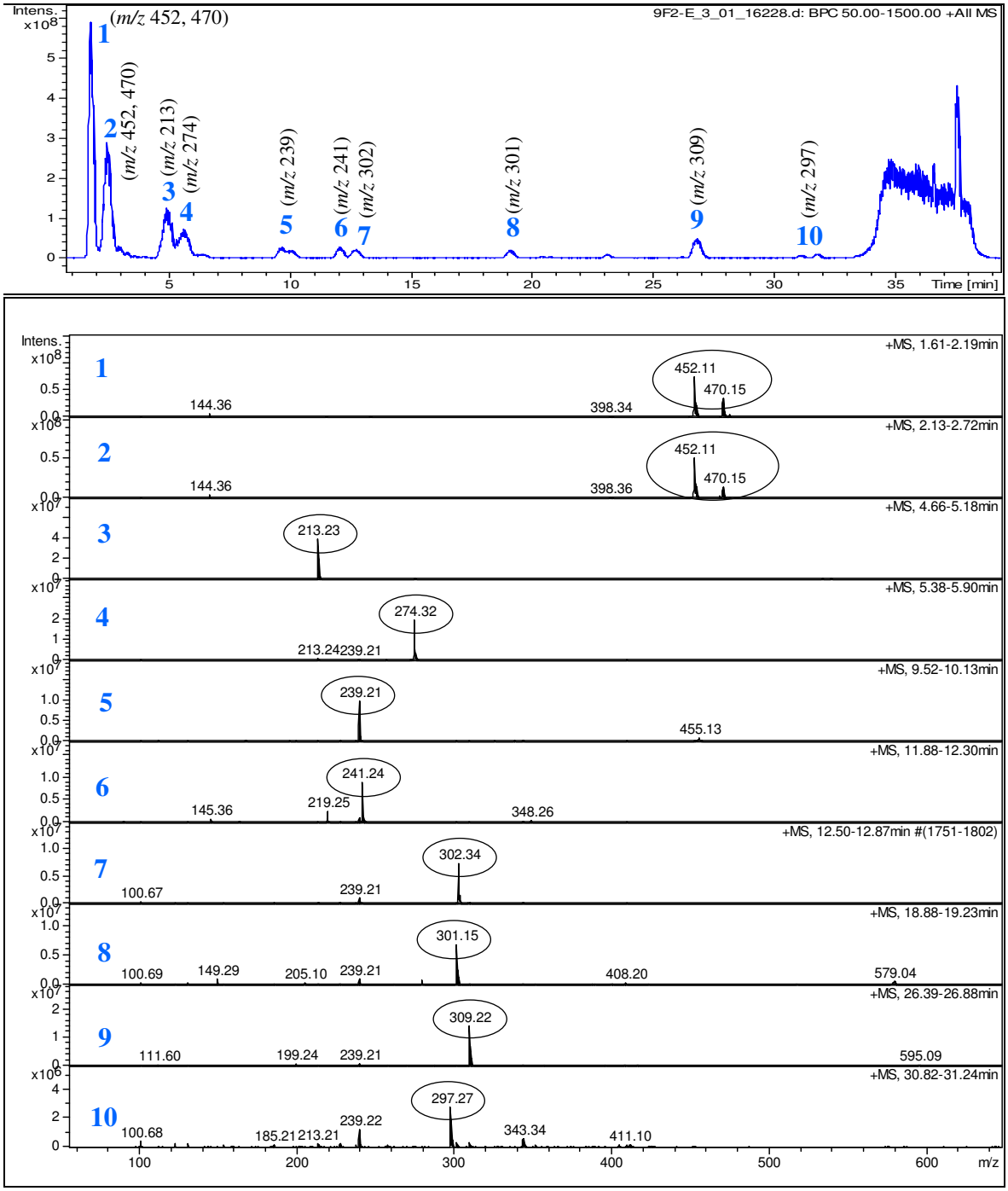


Figure 2-22 Chromatogram from LC-MS analysis of fraction 9F2E. Mass spectra for each of the numbered peaks are shown below the chromatogram and the major ion(s) in each spectrum are highlighted

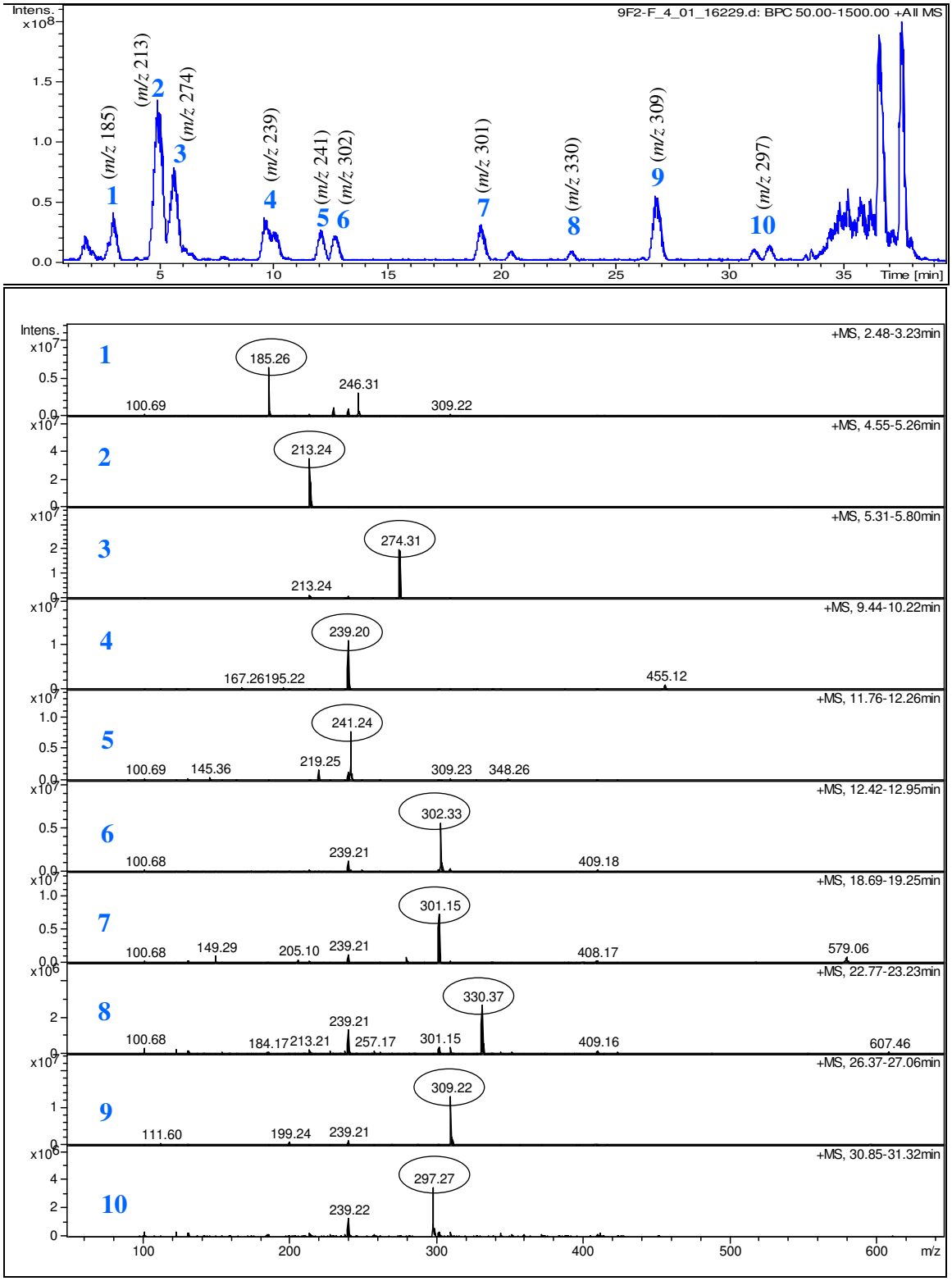


Figure 2-23 Chromatogram from LC-MS analysis of fraction 9F2F. Mass spectra for each of the numbered peaks are shown below the chromatogram and the major ion(s) in each spectrum are highlighted

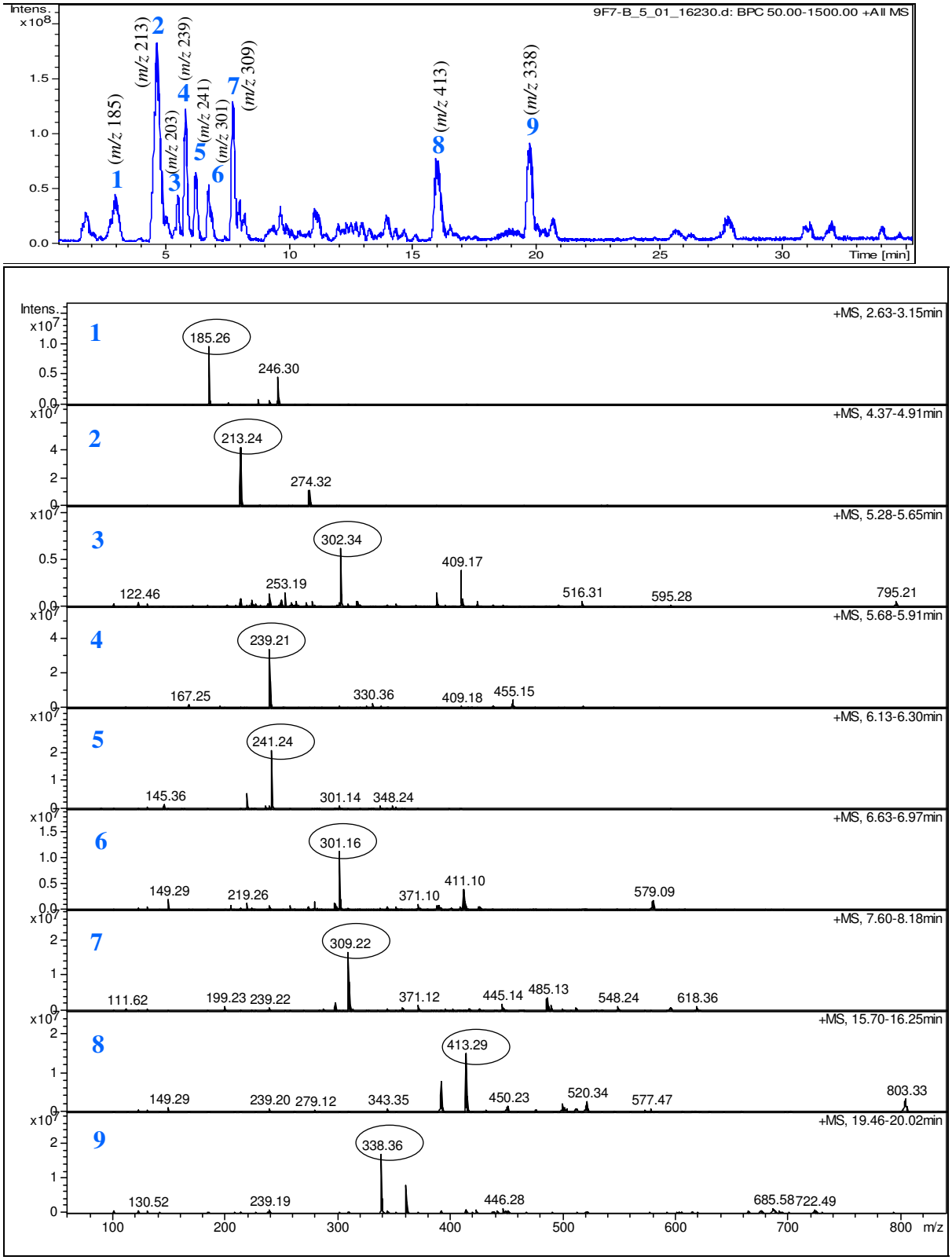


Figure 2-24 Chromatogram from LC-MS analysis of fraction 9F7B. Mass spectra for each of the numbered peaks are shown below the chromatogram and the major ion(s) in each spectrum are highlighted

Comparative analysis of the LC-MS chromatograms of the four most active fractions was carried out to establish the major metabolites common to all the fractions. The results are shown in graphical form in Figure 2-25.

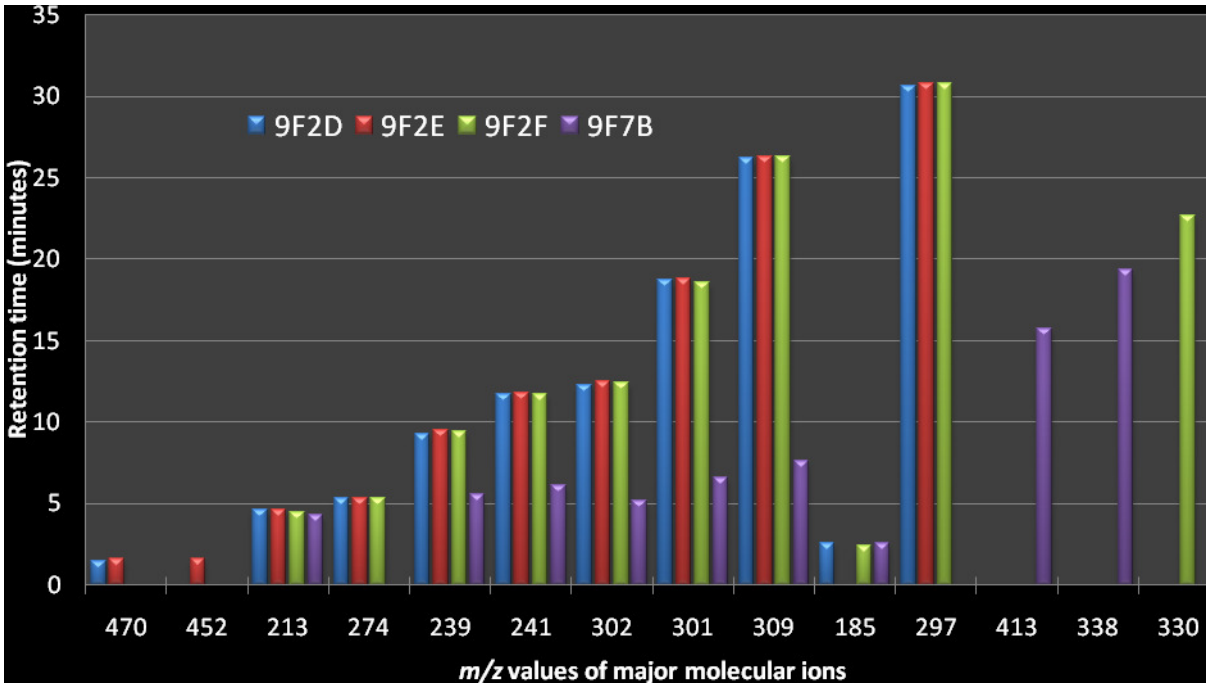


Figure 2-25 Graphical representation of the major metabolites associated with the major HPLC peaks within fractions 9F2D, 9F2E, 9F2F and 9F7B as determined by LC-MS analysis (Figures 2.21 – 2.24)

There were six major metabolite ions found common contained within the major peaks of the four peak fractions (9F2D, 9F2E, 9F2F and 9F7B) analysed on LC-MS chromatography. These metabolites are *m/z* = 213, 239, 241, 302, 301 and 309. In case of the peak fractions of the fraction 9F2 (9F2D, 9F2E and 9F2F) the retention times for these metabolite ions were fairly consistent but different in case of the fraction 9F7 (9F7B). For example the retention time of the compound with *m/z* = 309 was about 26 minutes for the peak fractions 9F2D, 9F2E and 9F2F but in the case of the peak fraction 9F7B the retention time for the counterpart compound with *m/z* = 309 was about 6.5 minutes. As a general rule of thumb, compounds giving similar *m/z* but different retention time are unlikely to be the same. However in this case there seems to be a pattern because there is more or less a persistent difference in retention times of the compounds with *m/z* 239, 241, 302 and 301 between the peak fraction 9F7B and the other 3 peak fractions, 9F2D, 9F2E and 9F2F. The difference in

retention time might be due to poor separation of the peak fraction sample, 9F7B, on LC column.

Based on the ichthyotoxic exposure results of the major peak fractions of fraction 9F2 and 9F7 (Table 2-4), peak fraction 9F2E caused the highest level of damage to fish gills. Analysis of the LC-MS chromatogram of the peak fraction 9F2E showed that there was one major peak and the major compound ion(s) contained within this peak were of *m/z* 470 and 452 (Figure 2-21 and 2-25). The compound with *m/z* 470 was also observed in case of fraction 9F2D. These molecular ions along with the compounds with *m/z* 301 and 309 which were common to all four peak fractions and also the major components were characterised by high resolution mass spectrometry (Table 2-5).

Table 2.5 Molecular formulae deduced from high resolution mass spectrometry analysis for major compounds in fractions 9F2D, 9F2E, 9F2F and 9F7B and a summary of the results of searches of the Chemical Abstracts database for compounds with the deduced molecular formulae (using SciFinder)

Meas. *m/z*	Ded. mol. formula	Err (ppm)	SciFinder search results
452.1722.	$C_{13}H_{30}N_3O_{14}$,	0.1	No hits
452.1704,	$C_{25}H_{26}NO_7$,	4.0	>400 hits
452.1717,	$C_{26}H_{22}N_5O_3$,	1.0	>500 hits (synthetic compounds including pesticides, anticancer, treatment of Parkinson's disease. HIV,
452.1736	$C_{14}H_{26}N_7O_{10}$	3.1	No hits
470.1828,	$C_{13}H_{32}N_3O_{15}$,	0.5	No hits
470.1809,	$C_{25}H_{28}NO_8$,	4.5	>200 hits (synthetic compounds including antidiabetic, antitumor, herbicide etc)
470.1823,	$C_{26}H_{24}N_5O_4$,	1.6	>700 hits (synthetic compounds, pesticides, anticancer etc)
470.1841	$C_{14}H_{28}N_7O_{11}$	2.3	No hits
301.1424	$C_{18}H_{20}O_4$	2.3	>3000 hits (synthetic compounds, pesticides, anticancer agents, insecticides, fungicides etc)
301.1408	$C_{14}H_{17}N_6O_2$	4.7	>400 hits (synthetic compounds including antiallergics, herbicides, insecticides, funcicidal, antibacterial etc)
301.1421	$C_{16}H_{19}N_3O_3$	2.3	> 40 hits (synthetic compounds)
301.1394	$C_{13}H_{21}N_2O_6$	5.9	>300 hits (herbicidal, antibacterial agents, antitumor agents etc)
309.2051	$C_{18}H_{29}O_4$,	3.9	>1500 hits (antitumor activity, artemisinin derivative (treatment of malaria, etc),
309.2034	$C_{14}H_{25}N_6O_2$	1.1	>100 hits (anti-microbial, herbicidal etc)

For each of the compounds, HRMS data were consistent with more than one possible molecular formulae. Most of the hits from the reference abstract database search (SciFinder) (Table 2-5) were of synthetic origin and were claimed to have a range of biological activities including anti-tumour, insecticidal, antibacterial, herbicidal, antidiabetic and fungicidal effects.

Major compounds identified in the 9F2E, 9F2F, 9F7B and 9F2D fractions were also analysed by MS/MS (Table 2-6). One of the major ions within fraction 9F2E had *m/z* = 452. This ion fragmented to give an ion with *m/z* = 144 which on further fragmentation suffered a neutral loss of 44 to give an *m/z* = 100 ion. Similarly, the other major parent ion with *m/z* 470 identified in this fraction fragmented to yield an ion with *m/z* =144 (Figure 2-26). The fragmentation pattern of the ions with *m/z* = 452 and *m/z* 470 seems to be the same as observed in the preliminary work (Section 2.4) carried out prior to the fish exposure experiments, in which some major metabolites within the Actino 13 culture supernatant were analysed. The MS/MS analyses also indicated that the compounds giving rise to the *m/z* = 470 and *m/z* = 452 ions are related (they both give the *m/z* = 144 daughter ion). Moreover the ion with *m/z* = 452 may have resulted from the *m/z* = 470 ion after dehydration (loss of water, MW = 18).

The compound giving rise to an ion with *m/z* 413 in the 9F7B fraction was also subjected to MS/MS analysis and produced daughter ions with *m/z* 357 and *m/z* 301. Further fragmentation of the *m/z* = 357 daughter ion produced a major ion with *m/z* = 238. The *m/z* 238 ion gave ions with *m/z* = 221 and 203 upon further fragmentation. The ion at *m/z* 413 could be a commonly observd background ion in mass spectrometry corresponding to a plasticiser (diisooctyl phthalate) which may appear at *m/z* 391 (M+H) and *m/z* 413 (M+Na).

Table 2.6 Major compounds identified in the fractions 9F2E and 9F7B on MS/MS and their major fragment ions

Parent ions (*m/z*)	Major daughter ions (*m/z*)
452	144 (100)
470	144
413	357(238), 301, 238(221,203)

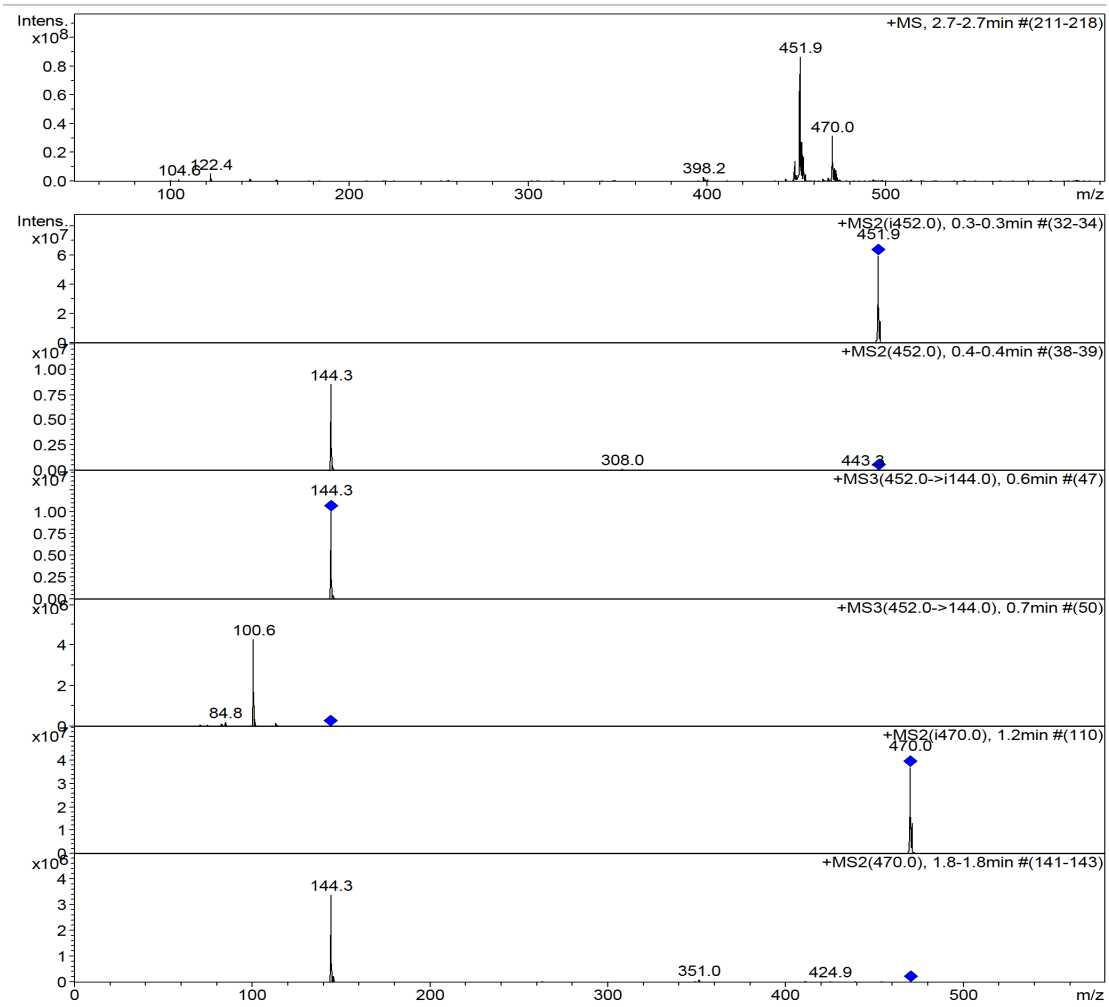


Figure 2-26 MS-MS spectra of ions with *m/z* 452, and *m/z* 470 in fraction 9F2E

2.6 Conclusions and future work

The aim of this project was to investigate fish toxins produced by *Streptomyces* species. Due to the limited availability of fish, exposures of fish toxins-containing samples could only be carried out twice a year.

Samples for fish-exposure experiments were prepared from an actinomycete belonging to the *S. griseus* clade isolated from a site of major fish kills, as well as from the type strain, *Streptomyces griseus* DSM 40236. It was shown that the type strain produced ichthyotoxic metabolites as well as the environmental strain. This is a novel and unexpected observation.

The evolution of the strategy for sample preparation and separation was based on the results of each round of fish exposure experiments. After a series of consecutive HPLC separations, we were able to narrow down the highest level of fish toxicity to four fractions. These four fractions each contained only a handful of compounds. Some of these compounds were purified and analysed by high resolution mass spectrometry to determine their molecular formulae.

Comparison of the metabolic profile of these most toxic fractions showed that many of the compounds were common to all of them e.g. compounds yielding ions with *m/z* 213, 241, 239, 301 and 309 were present in all the fractions. A compound giving an ion with *m/z* = 470 was the major compound in the two most active fractions i.e. 9F2E and 9F2D which showed 90.3% and 73.3% fusion of secondary lamellae, respectively, after 96 hours of exposure. Another compound associated with the most active fraction 9F2E gave an ion with *m/z* = 452. This was not observed in the LC-MS chromatograms of any other fractions. HRMS data indicates that the compounds with *m/z* = 470 and 452 might be related and it is suggested that *m/z* = 452 results from *m/z* = 470 by dehydration. The molecular formulae determined for the compounds giving rise to the *m/z* = 452 and *m/z* = 470 ions were used to search the chemical abstracts database (using SciFinder). One of the possible molecular formulae for each of the compounds with *m/z* = 452 and 470 returned 100s of hits while, the other two possible molecular formulae for these compounds returned no hits, suggesting the compounds might be novel.

The sample preparation and purification route for the fish toxins has been fully established in this work. Thus, future experiments could involve purification of the probable fish toxins e.g. the compounds giving rise to ions with *m/z* = 452 and 470 followed by fish-exposure experiments with the purified compounds.

Although ichthyotoxicity was found to be very high (>90% of fusion of secondary lamella on fish gills after 96 hours exposure) in fraction 9F2E, the fact that similar damage on fish gills was observed in other fractions (such as 9F2D, 9F2E, 9F7B) to a slightly lesser extent, strongly suggests that there is more than one ichthyotoxic metabolite being produced by this strain. Also, the finding that the ichthyotoxic

metabolites are being produced by the type strain *Streptomyces griseus* DSM 40236 is a novel observation and very interesting for future research on the subject.

Due to the limited availability of fish for exposure experiments we were not able to make sufficient progress to reach to a firm conclusion. However, methods and protocols have been established that will be useful for future experiments

Part II: Structural and Biosynthetic Investigations of Foroxymithine

CHAPTER 3 - INTRODUCTION

3.1 Foroxymithine

Foroxymithine was first isolated in 1985 by Umezawa *et al.*, in the course of testing the ability of culture filtrates of freshly-isolated actinomycete strains to inhibit angiotensin-converting enzyme (ACE) [55]. The strain responsible for production of foroxymithine was taxonomically identified as *Streptomyces nitrosporeus* [55]. Foroxymithine is well known for its pharmacological activity as a ferric iron chelator that potentiates the antineoplastic agent erbstatin, and as an inhibitor of angiotensin-converting enzyme (ACE) and other proteases [176]. Conceivably, foroxymithine could function as a microbial siderophore (involved in ferric iron uptake), but to date no investigation of this hypothesis has been published.

The structure of foroxymithine (Figure 3-1) consists of four amino acid residues including one serine, and one residue each of *N*-α-acetyl-*N*-δ-hydroxy-*N*-δ-formylornithine, *N*-δ-hydroxyornithine and *N*-δ-hydroxy-*N*-δ-formylornithine

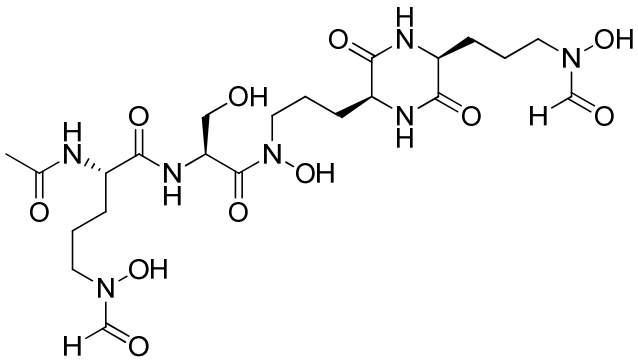


Figure 3-1 Structure of foroxymithine

Recently, Dr Shinya Kodani, a visitor to the Challis lab, identified foroxymithine as a metabolite of *S. narbonensis*. There is scant literature on the secondary metabolites of *S. narbonensis*. It is known to produce josamycin, a macrolide antibiotic that shows strong activity against Gram-positive bacteria [177, 178] (Figure 3-2). Other known metabolites of *S. narbonensis* include the 14 membered macrolide antibiotics narbomycin and picromycin [179] (Figure 3-2)

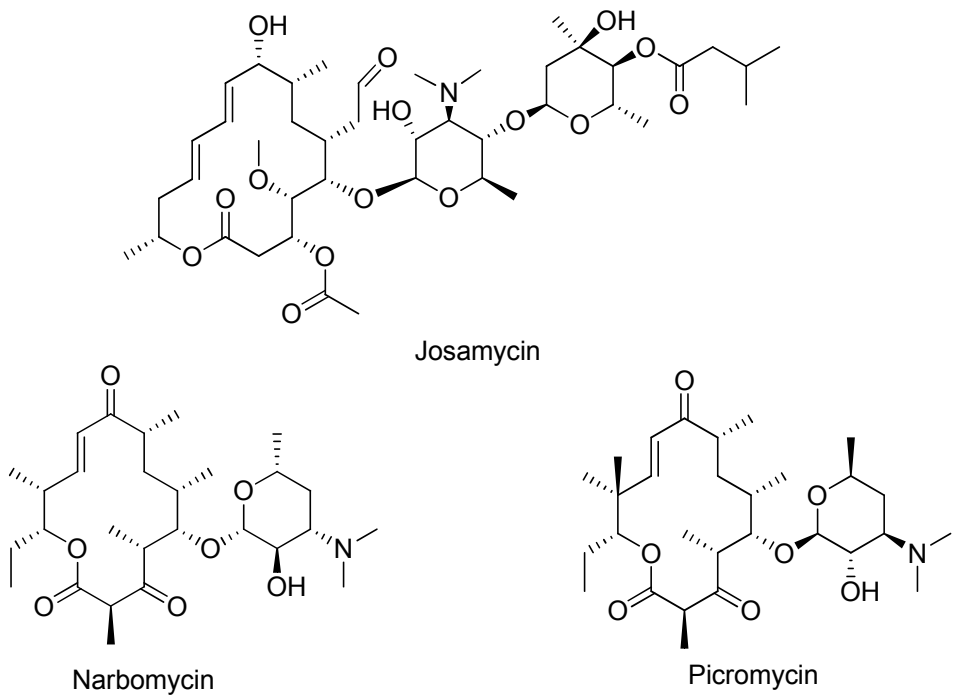


Figure 3-2 Chemical structures of josamycin, narbomycin and picromycin

3.2 Structures of foroxymithine and coelichelin compared

The structure of foroxymithine, originally isolated from *Streptomyces nitrosporeus* was proposed to contain all L-stereochemistry at the four α-carbon atoms [55]. The structure has remarkable similarity to a known siderophore, coelichelin (Figure 3-3) which was first isolated from *Streptomyces coelicolor* A3 (2) by the Challis group using a genome mining approach [180].

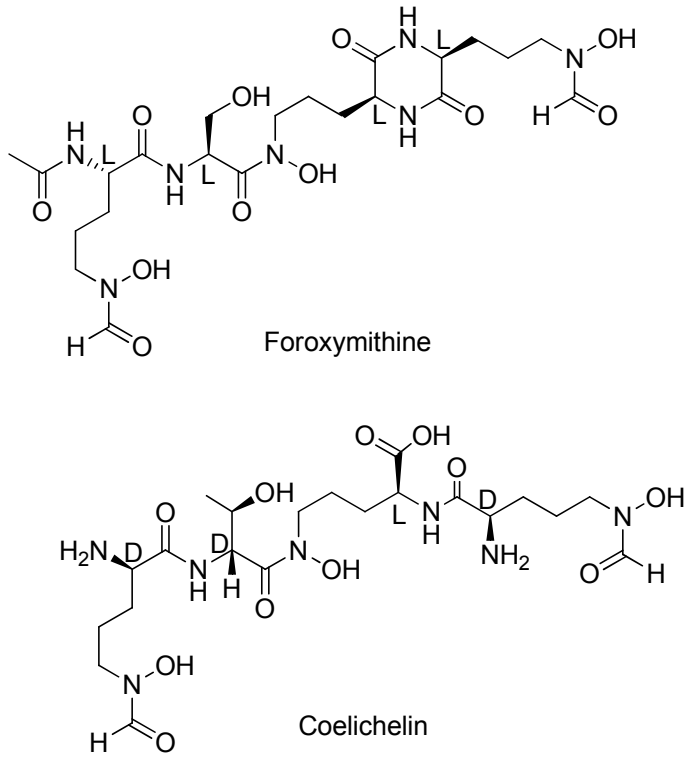


Figure 3-3 Comparison of chemical structures of foroxymithine and coelichelin

Coelichelin is a tetrapeptide consisting of a threonine residue, two residues of *N*-δ-hydroxy-*N*-δ-formylornithine and one residue of *N*-δ-hydroxyornithine. Coelichelin and foroxymithine are remarkably similar in structure, which suggests that they might share similar biosynthetic pathways. However, the relative stereochemistry at the four α-carbon atoms, of coelichelin, has been proposed to be D,D,L,D, on the basis of a combination of 1D and 2D NMR spectroscopy experiments and molecular modelling [180]. The difference in the stereochemistry of coelichelin and foroxymithine is very surprising given that they have such similar planar structures. Proton NMR data for foroxymithine isolated from *S. nitrosporeus* was reported in a 1985 patent application [181] and the total synthesis of foroxymithine was reported in 1991 [176]. However the ^{1}H NMR signals reported for synthetic and natural foroxymithine are very different. No experimental data to support the all L-stereochemical assignment of natural foroxymithine has been published and only the relative stereochemistry of coelichelin has been experimentally determined. The D, D, L, D absolute stereochemistry for coelichelin was hypothesised on the basis of biosynthetic considerations, but requires experimental verification.

3.3 Non-ribosomal peptide biosynthesis

A range of biologically-active peptides are produced by a variety of microorganisms including bacteria and fungi. These molecules are generally produced under special physiological conditions [182]. Structural features such as non-proteinogenic amino acids, (including D-amino acids), hydroxy acids and fatty acids, small heterocyclic rings and other unusual modifications in the peptide backbone are common in these molecules. It has been established that in many cases, ribosomes, transfer RNA (tRNA) and messenger RNA (mRNA) are not required for the formation of these peptides [183, 184]. It was about 30 years ago when Fritz Lipmann first proposed that the multi-enzyme peptide synthetases employ the thiotemplate mechanism in the biosynthesis of such peptides [185, 186].

Compounds that are biosynthesized by nonribosomal peptide synthetases (NRPSs) possess extraordinary pharmacological importance. Antibiotics e.g. penicillin [187-189], vancomycin [190-192], bacitracin [193] and gramicidin [194], immunosuppressive agents such as cyclosporin [189, 195] and siderophores such as mycobactin [196], enterobactin [197] and coelichelin [83, 198] are all assembled by NRPS dependent mechanisms. The structures of some of these important molecules are shown in Figure 3-4.

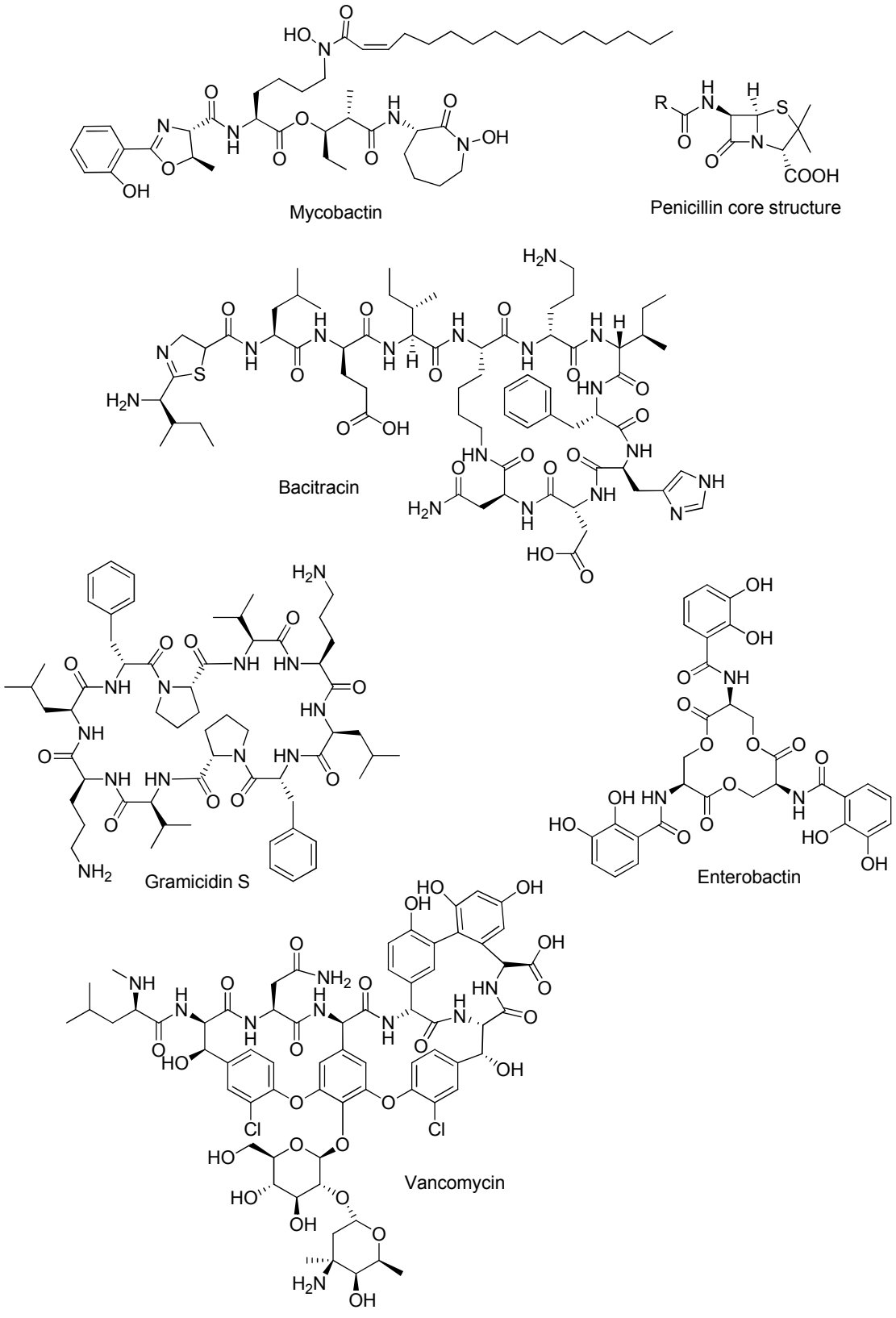


Figure 3-4 Chemical structures of bioactive compounds assembled by nonribosomal peptide synthetases

NRPSs are large multifunctional enzymes ranging in size from 100-1700KDa [199]. Each module within an NRPS is responsible for introducing one amino acid residue into the final product. NRPS modules can be further subdivided into domains which are individual enzymes that perform specific functions [200].

In every NRPS chain extension module, there are three domains which are always present. The first is the adenylation (A) domain, which is responsible for specifically recognising a substrate amino (or hydroxy) acid and activating it as an amino-acyl adenylate by reaction with ATP. The second is the peptidyl carrier protein (PCP) domain (also known as the thiolation (T) domain) to which substrate residues, activated by the A domain, are attached as thioesters. The third is the condensation (C) domain which is responsible for catalysing peptide bond formation between aminoacyl thioesters attached to PCP domains within adjacent modules. These are the core domains of every chain extension module, but other domains can also be present, such as epimerisation (E) and N-methyl transferase (MT) domains. There is usually a thioesterase (TE) domain at the C-terminus of the last module that catalyses release of the fully assembled peptide chain via cyclisation or hydrolysis.

A crystal structure of the last NRPS module (SrfA-C) of the NRPS responsible for surfactin biosynthesis in *Bacillus subtilis* was reported in 2008 [201]. The module contains C, A, PCP, and TE domains. SrfA-C catalyses the addition of $_{L}$-leucine to a hexapeptidyl thioester intermediate in surfactin biosynthesis, as well as the release of the lipopeptide surfactin by cyclisation. The reported 2.6 angstrom crystal structure (Figure 3-5) provides insight into how the catalytic domains coordinate with each other within an NRPS module. The structure shows close association of adenylation and condensation domains to form a catalytic platform, with their active sites on the same side of the platform. The attachment of the PCP domain to this platform is very flexible. Location of the PCP domain is ideal because once its 4'-phosphopantetheinyl (4' Ppant) arm is loaded with a substrate; it can easily reach to the active site of the adenylation domain and donar site of the condensation domain. The crystal structure provides the first structural understanding of the design of an entire NRPS module. It shows how different conformational changes in the PCP domain enables it to reach to every essential catalytic domain [202].

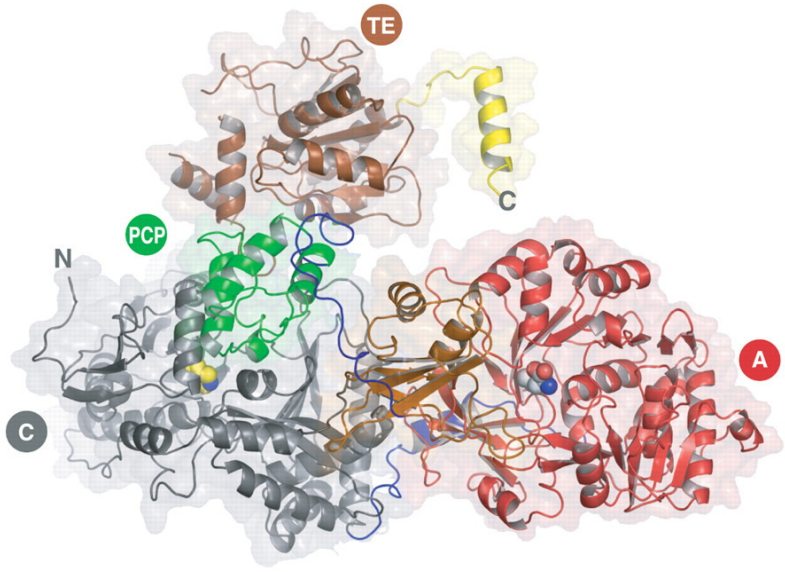


Figure 3-5 Overall structure of the last NRPS module responsible for assembly of surfactin (SrfA-C) in *B. Subtilis*. The C domain's (gray) catalytically active residue His147 and a leucine residue bound in the A domain's (red and orange) active site are shown in space-filling representatin along with PCP (green) and TE (brown) domains connected via short linker regions [201]

3.3.1 The catalytic mechanism and substrate specificity of NRPS adenylation domains

The adenylation (A) domain is a key component of NRPSs. It is approximately 550 amino acids in length [203]. This domain performs an indispensable role in selection and activation of the amino acid residues to be incorporated into the growing peptide chain. In order to activate an amino acid residue, a two-step mechanism is employed. In the first step, the activation of the carbonyl group of the amino acid takes place by making an aminoacyl adenylate using ATP and Mg^{2+}. In the second step the aminoacyl adenylate is attached to the 4'-phosphopantetheinyl (4' Ppant) cofactor of the PCP domain (also called the thiolation (T) domain) by a thioester linkage (Figure 3-6). This cofactor is post-translationally attached to a conserved serine residue of PCP domains by a phosphopantetheinyl transferase [204]. The PCP domain represents the transport unit, which enables the activated amino acids and elongation intermediates to move between the catalytic centres. Any modifications, such as epimerisation or methylation, take place while the aminoacyl thioester is still loaded onto the PCP domain. This is followed by chain elongation catalysed by the condensation (C) domain.

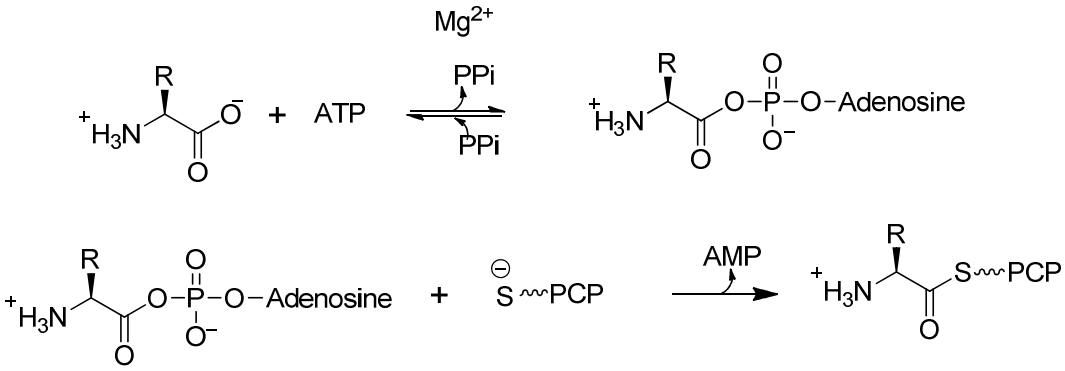


Figure 3-6 Reaction scheme showing activation of an amino acid by an A domain as an aminoacyl adenylate followed by its transfer on to a PCP domain to form an aminoacyl thioester

A domains are usually very specific in selecting and activating an amino or hydroxy acid. However, specificity is principally but not exclusively controlled by the A-domain. Condensation domains have also been reported to show some substrate selectively but this has yet to be fully elucidated experimentally [205]. The specificity of A domains has been studied extensively in the last few years. Conti *et al*. were the first to successfully crystallise an A domain. This was from the gramicidin S synthetase GrsA and was co-crystallised with the amino acid L-phenylalanine and AMP (Figure 3-7) [206]. Each A domain comprises a large N-terminal core structural domain (~450 amino acids long) shown in blue and a small C-terminal structural subdomain (~100 amino acids long) shown in green. The two domains are connected by a small 5-10 amino acid residue hinge. The core and subdomain structural organisation of A domains is highly conserved [202]. The C-terminal structural subdomain can adopt multiple orientations relative to the N-terminal structural domain during the catalytic cycle of the A domain. The cycle starts with the open conformation in which the C-terminal subdomain bends away from the active site which promotes entry of substrate amino acids. In the next step adenylation of the substrate amino acid residue takes place in the closed state and finally the subdomain lifts up (open state) again to transfer the aminoacyl moiety to the reactive thiol group of the 4'-phosphopantetheine arm of the downstream PCP domain [201, 202].

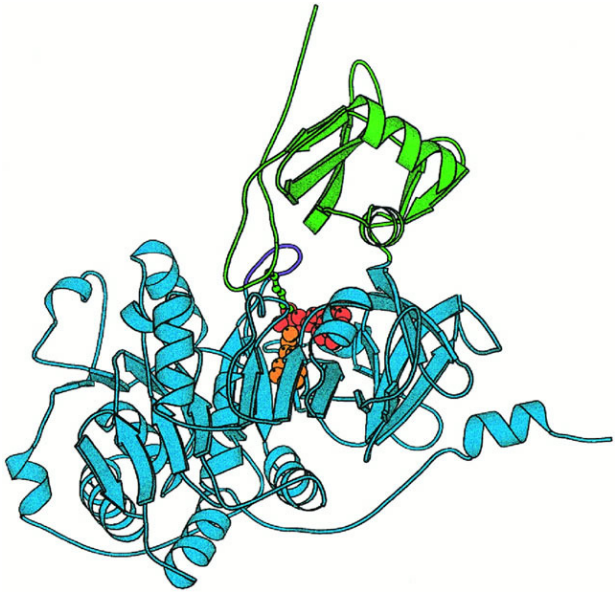


Figure 3-7 Ribbon diagram of the phenylalanine activating domain of gramicidin synthetase 1 (PheA) in a ternary complex with phenylalanine and AMP based on its crystal structure: Blue indicates the large N-terminal domain and green represents the small C-terminal domain [206]

This crystal structure allowed the identification of the L-phenylalanine binding pocket and the residues lining the pocket. In another study, Challis *et al.,* analysed the protein sequences of 150 NRPS activation domains in comparison to phenylalanine-selective A domain of GrsA, to pinpoint the essential amino acid residues involved in substrate specificity and binding. It was suggested that eight residues lining the binding pocket were sufficient to both predict and, in many cases, rationalize the substrate specificity of NRPS A domains. A phylogenetic tree was constructed on the basis of amino acid sequence alignments of A domains, potentially a useful tool for predicting the order and identity of amino acids activated by a nonribosomal peptide synthetase on the basis of protein sequence data [198]. Marahiel and co-workers also analysed the binding pocket of phenylalanine specific A domain in comparison with the corresponding moieties in other A domains and hypothesised that there are 10 residues involved in binding the cognate substrate amino acid [184, 207]. Two conserved residues, Asp235 and Lys517 were identified as responsible for electrostatic interaction with the α-amino and α-carboxylate groups of the substrate. The remaining eight residues, Ala236, Trp239, Thr278, Ile299, Ala301, Ala322, Ile330 and Cys331, interact with the side chain [207]. Of the remaining eight residues, three residues (Ala236, Ala301 and Ile330) are only moderately variable. Three other residues (Trp239, Ala322 and Cys331) were found to be highly variable, which suggests that they may have a major role in A

domain substrate selectivity. Two additional residues (Thr278 and Ile299) are also highly variable but they were not consistently found to be involved in A domain substrate selectivity. The structure of a second A domain was solved in 2002 by May *et al.* This showed a very similar overall fold [208]. It allowed the extension of the empirical correlations between protein sequence and substrate specificity to some aryl acid activating A domains. The first structure of a eukaryotic A domain from the siderophore-synthesizing NRPS, SidN, from the endophytic fungus *Neotyphodium lolii* was reported by Lee *et al.,* in 2010 [209]. By sequence analysis of the A domains of a biochemically uncharacterised NRPS, it is now possible in many cases to predict the substrates it utilises.

The discovery of the siderophore coelichelin in *S. coelicolor* is a pioneering example of how A domain comparative sequence data was used to predict the amino acid residues incorporated into a novel nonribosomal peptide natural product [83]. In a similar study, Bruner and co-workers reported the discovery of another new peptide siderophore fuscachelin (Figure 3-8) from the moderate thermophile *Thermobifida fusca* [210]. Since the first reported example of accurate A domain specificity prediction by the Challis group in 2000, several other examples have appeared in the literature. These include the discovery of erythrochelin in *Saccharopolyspora erythraea*, and antichelin from *Streptomyces scabies* [211].

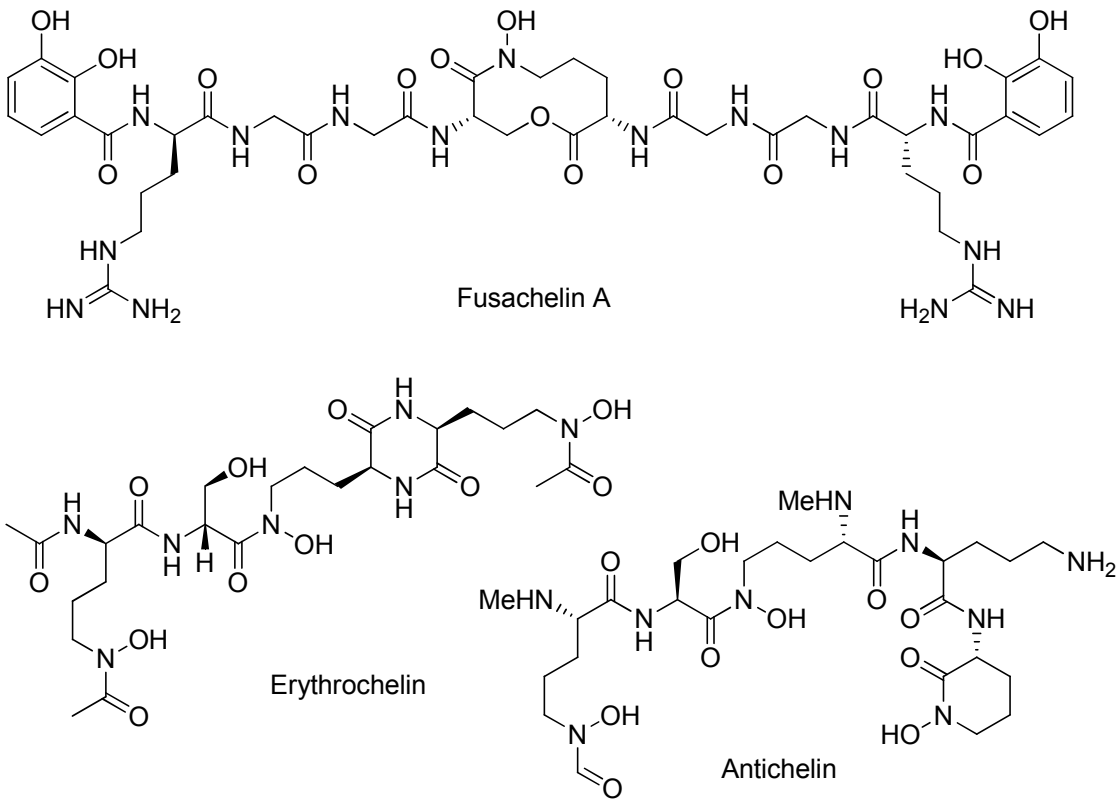


Figure 3-8 Chemical structures of novel nonribosomal peptides discovered via prediction of NRPS A domain substrate specificity

3.3.2 Role of PCP domains

Peptidyl carrier protein (PCP) or thiolation (T) domains usually contain 80-100 amino acids and are located directly downstream of A domains [212]. Amino acids are first selected and activated by the A domain and then transferred onto the PCP domain, which is responsible for the transfer of the attached aminoacyl or peptidyl thioesters to catalytic centres (such as condensation or epimerisation domains) [212]. A solution structure of the PCP domain of the *Bacillus brevis* tyrocidine synthetase 3 (TycC), solved using NMR techniques, is shown in Figure 3-9 [213]. It shows a distorted four-helix bundle with an extended loop between the first two helices. Overall the topology is somewhat similar to the acyl carrier proteins (ACPs) from the *Escherichia coli* fatty acid synthase and the actinorhodin polyketide synthase from *S. coelicolor.* The conserved serine residue, which is the binding site for the Ppant cofactor, has the same location as in the ACPs. A crystal structure of a PCP-C didomain from the tyrocidine synthetase TycC has also been reported (Figure 3-11). The PCP domain

within this didomain fragment has a similar structure to the isolated PCP domain of TycC [202, 214].

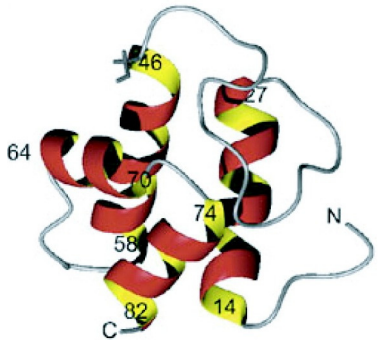


Figure 3-9 Ribbon diagram of the PCP domain of the *Bacillus brevis* tyrocidine synthetase 3 (TycC) solved using NMR techniques. Residue numbers are located at the first and last amino acid of each helix [213]

3.3.3 Catalytic mechanism of peptide chain elongation by condensation domains

Condensation (C) domains are responsible for peptide bond formation. They catalyse nucleophilic attack of the α-amino group of the PCP-bound downstream amino acid residue on the thioester carbonyl group of the upstream PCP-bound amino acid residue (Figure 3-10) [215]. C domains contain ~450 amino acids. The structure of the C domain VibH was reported by Keating *et al*. This provided insight into the structural basis for the catalytic function of this domain [216]. VibH catalyses the condensation of a PCP-bound 2,3-dihydroxybenzoyl (DHB) residue with norspermidine [217]. VibH differs from most other C domains because of the fact that norspermidine is not covalently attached to a PCP domain but is free in solution. The crystal structure by Keating *et al*., revealed the existence of a solvent exposed channel into which the DHB-Ppant chain could bind.

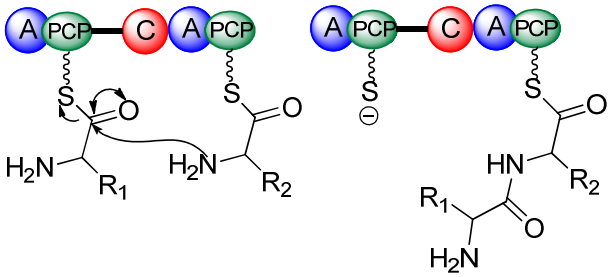


Figure 3-10 The proposed catalytic mechanism of condensation (C) domains

C domains have been reported to possess significant substrate selectivity towards both the absolute stereochemistry of the α-carbon and the nature of the amino acid side chain [215, 218]. This observation suggests the presence of aminoacyl thioester binding pockets within C domains [218]. This may explain why peptide chain assembly initiates within the first module of the NRPS and not within other chain elongation modules. This ultimately controls the directionality of nonribosomal peptide biosynthesis [219].

Crystal structures of several NRPS C domains have been reported. These include the free standing C domain (VibH) from the vibriobactin synthetase of *Vibrio cholerae* [216] and the C domains from within the enterobactin (EntF) [220] and tyrocidine (TycC) synthetases [214]. All of these structures have revealed a V-shaped architecture for C domains. Figure 3-11 shows a crystal structure of the PCP-C didomain of the tyrocidine synthetase TycC. Within the V-shaped structure the two equally big structural subdomains are connected only by a small hinge region at the bottom of the V and by a bridging loop in the middle, which results in a canyon-like active site structure with the catalytically active H224 (shown in yellow) in the middle. Thus, the C domain is accessible to the substrates-loaded Ppant arms of the corresponding upstream or downstream PCP domains . [201, 202, 214]

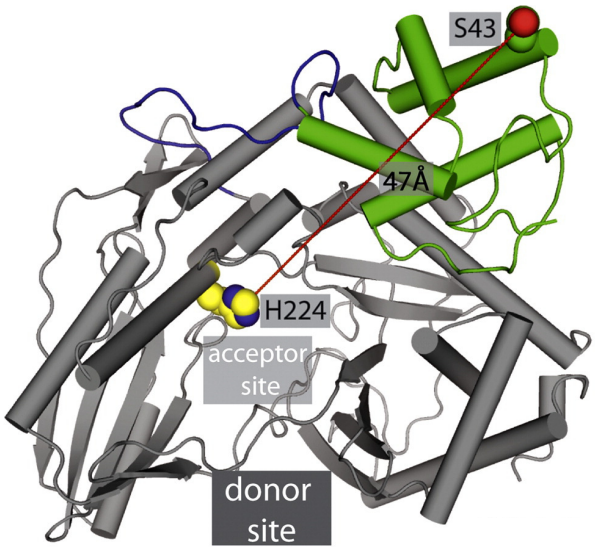


Figure 3-11 X-ray structure of PCP-C didomain of the tyrocidine NRPS TycC. The PCP domain is shown in green with S43, the site of posttranslational phosphopantetheinylation, highlighted. The C domain is shown in grey with the active site histidine 224 residue highlighted as yellow/blue spheres [214]

3.3.4 Role of epimerisation domains

A large number of peptides biosynthesised by NRPSs contain D-amino acids. In principle, there are two ways to introduce a D-amino acid into the backbone of a peptide through NRPS machinery. There are D-amino acid specific A domains in some NRPSs which specifically select and activate D-amino acids. One example is cyclosporin synthetase, which contains an A domain that is specific for D-alanine [221]. It is also suggested that the D-Glu residue in microcystin is introduced by an A domain that is specific for D-amino acids. The second and most common way to introduce D-amino acids into nonribosomal peptides is via epimerisation (E) domains, which contain about 450 amino acids, exhibit sequence similarity with C domains, and are located (when present) downstream of PCP domains [199, 221]. E domains represent a class of cofactor independent amino acid epimerases that catalyse the de- and re-protonation of the α-carbon atom of an enzyme bound aminoacyl or peptidyl-thioester. They can catalyse epimerisation in both directions i.e. L to D and D to L [222].

Epimerisation involves cleavage of the Cα-H bond of an aminoacyl thioester to yield the Cα-carbanion as a planar intermediate that can be re-protonated to form either the L or D product. It has been proposed that this reaction can take place either via a two-base mechanism, where one enzymatic base removes the Cα-H from the substrate and the conjugate acid of a second enzymatic base provides a proton to the opposite face, or via a one-base mechanism, where a single enzymatic base is responsible for both removal and addition of a proton [221]. These alternative mechanisms are shown in (Figure 3-12).

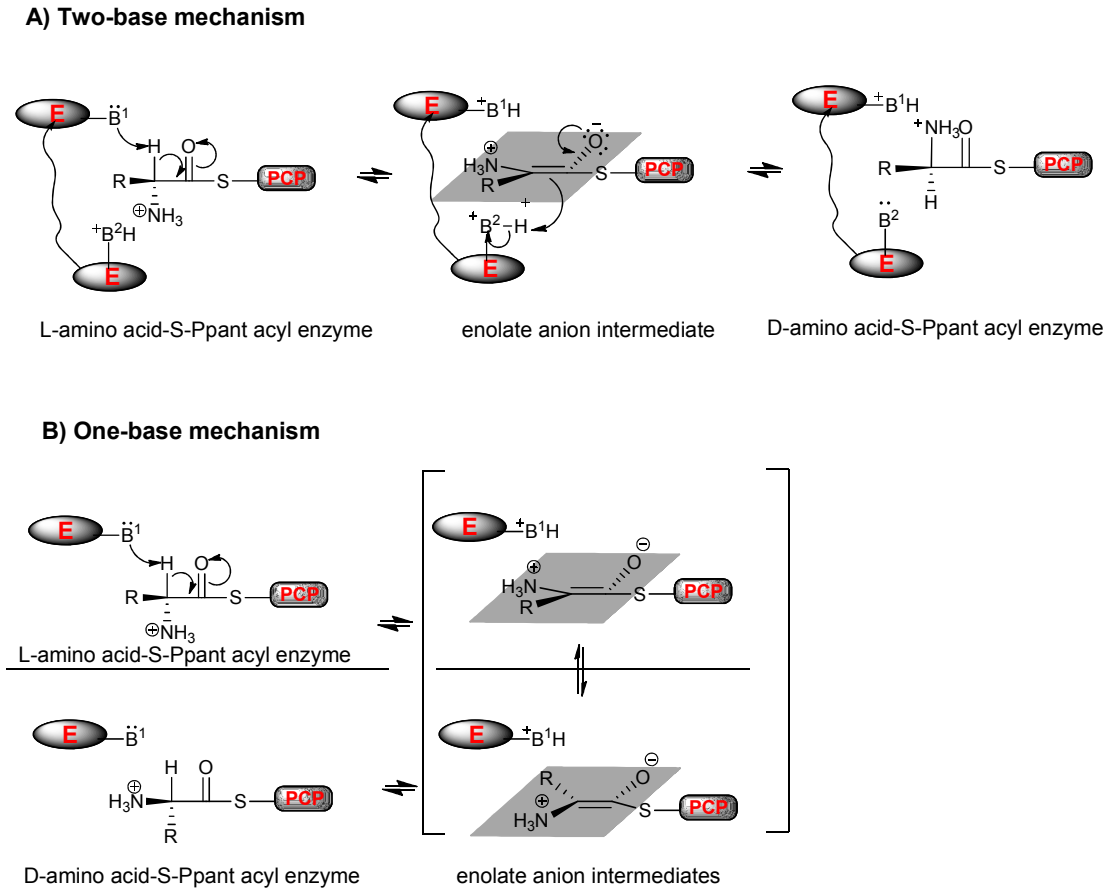


Figure 3-12 Alternative mechanisms to describe Cα-proton transfer during aminoacyl-S-Ppant epimerisation

Isolated E domains create a D/L mixture *in vitro*. Clugston *et al*. have shown that *in vivo* the selection of the correct stereoisomer from this mixture is mediated by the enantioselective donor site of the downstream C domain [223].

In an experiment to understand the basis for the specific interaction of E domains with PCP domains and to investigate their substrate tolerance, a set of eight fusion proteins was constructed. E and PCP-E domains of tyrocidine synthetase A (TycA) were fused to different gene fragments encoding A and A-PCP domains [224]. This experiment showed that the E domain of TycA, which usually epimerises L-Phe, also accepts alternate substrates such as Trp, Ile and Val, with reduced efficiency. However, only fusion proteins where the PCP domain originates from modules containing an E domain showed epimerisation activity. Sequence comparisons of such PCP domains revealed that epimerisation activity is influenced by the nature of amino acid residues

in proximity to the phosphopantetheine covalent attachment site. The aspartate residue in front of the conserved serine residue to which phosphopantetheine is attached was reported to play an important role for the mutual functioning of PCP and E domains. This strongly suggests that specialised PCP domains are needed for productive interactions with E domains.

The portabilities and substrate specificities of aminoacyl-E domains (E domains from initiation modules) and peptidyl-E domains (E domains from elongation modules) was investigated. Both aminoacyl- and peptidyl-E domains were able to epimerise both aminoacyl- and peptidyl-thioester substrates [225]. In a study of recombinant bimodular proteins derived from the tyrocidine synthetase B, the effect of exchanging peptidyl- and aminoacyl-E domains on enzymatic dipeptide formation was investigated and it was suggested that the epimerisation of an aminoacyl-S-Ppant substrate is more challenging than that of a peptidyl-S-Ppant substrate [225].

3.3.5 Peptide release mechanisms in NRPSs

After all the steps of chain elongation, the fully assembled peptide chain is attached as a thioester to the phosphopantetheine arm of the PCP domain of the last module of the NRPS. The thioester needs to be cleaved to make the multienzyme complex available for the next cycle of peptide formation. In most cases a thioesterase (TE) domain at the C-terminus of the NRPS is responsible for thioester cleavage. TE domains contain about 280 amino acids. All TE domains have a conserved catalytic triad, Ser-His-Asp, involved in release of the product. The His residue is stabilised by the adjacent Asp residue and acts as a catalytic base to remove a proton from the hydroxyl group of the Ser residue. Concomitant with proton abstraction, the hydroxyl group of the Ser residue attacks the carbonyl carbon of the peptidyl thioester bound to the adjacent PCP domain resulting in formation of a TE-bound oxoester. The TE-bound oxoester can be attacked at the carbonyl carbon by an external nucleophile (typically water) resulting in a linear peptide product (Figure 3-13). This mechanism is seen in the biosynthesis of yersiniabactin and vancomycin. Alternatively, the carbonyl carbon of the oxo-ester can undergo intramolecular attack by a hydroxyl or amino group within the assembled peptide chain resulting in a cyclic or branched cyclic product. This type

of mechanism is seen in the biosynthesis of a large number of nonribosomal peptides, including tyrocidine A and daptomycin [212, 226-228].

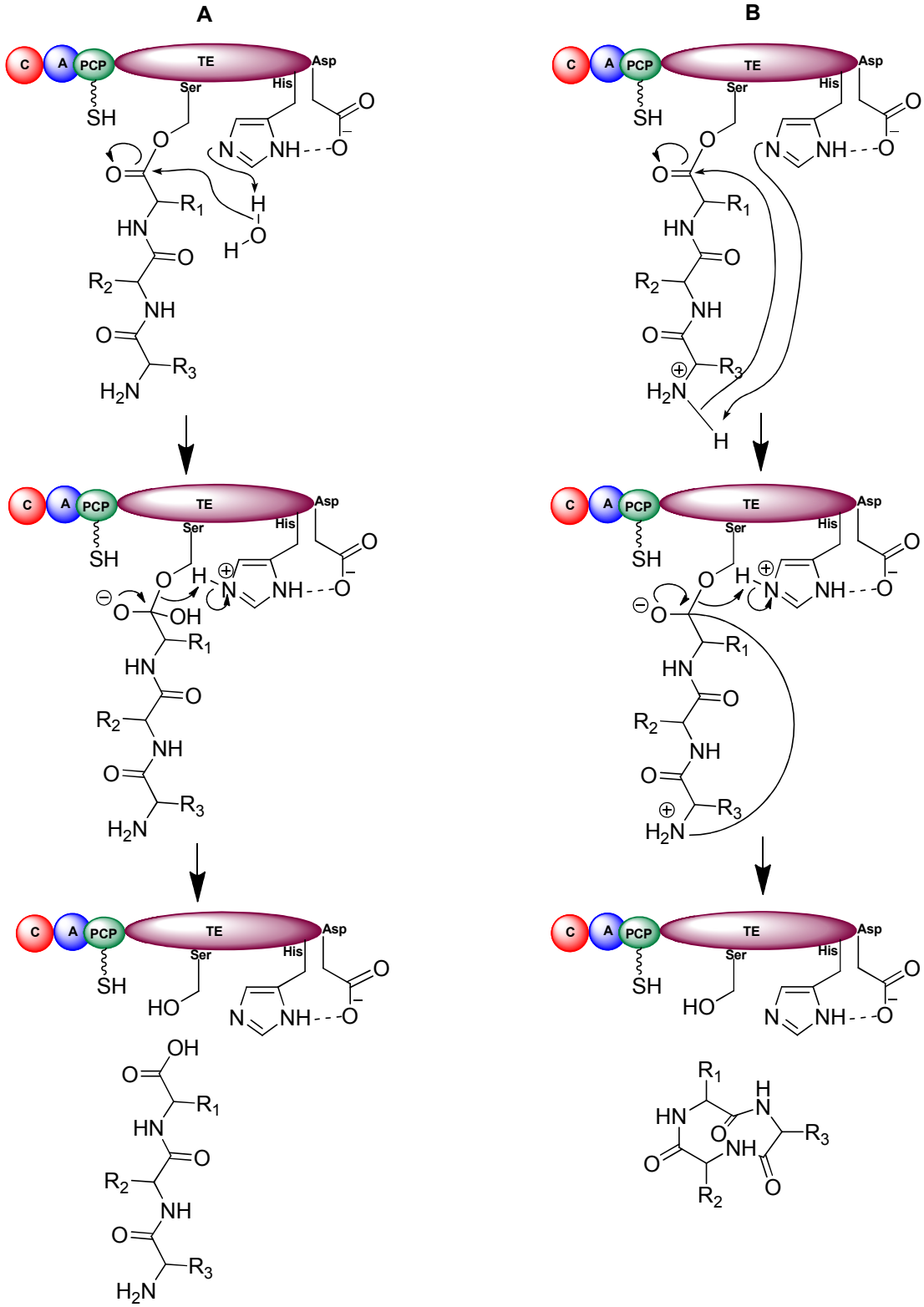


Figure 3-13 Mechanisms of peptide release reactions catalysed by TE domains. A) release of a linear peptide by hydrolysis and B) release of a cyclic peptide by intramolecular nucleophilic attack of the N-terminal amino group on the acyl-enzyme intermediate.

An alternative mechanism of chain release occurs by the action of iterative TE domains [227] . For example, in the biosynthesis of the symmetrical decapeptide gramicidin S, two identical pentapeptides are biosynthesised and then linked together. During assembly, the first pentapeptide is loaded onto the TE domain and when the second molecule of pentapeptide is assembled, the TE domain initially catalyses dimerisation and later cyclisation to release the final decapeptide gramicidin S (Figure 3-14)

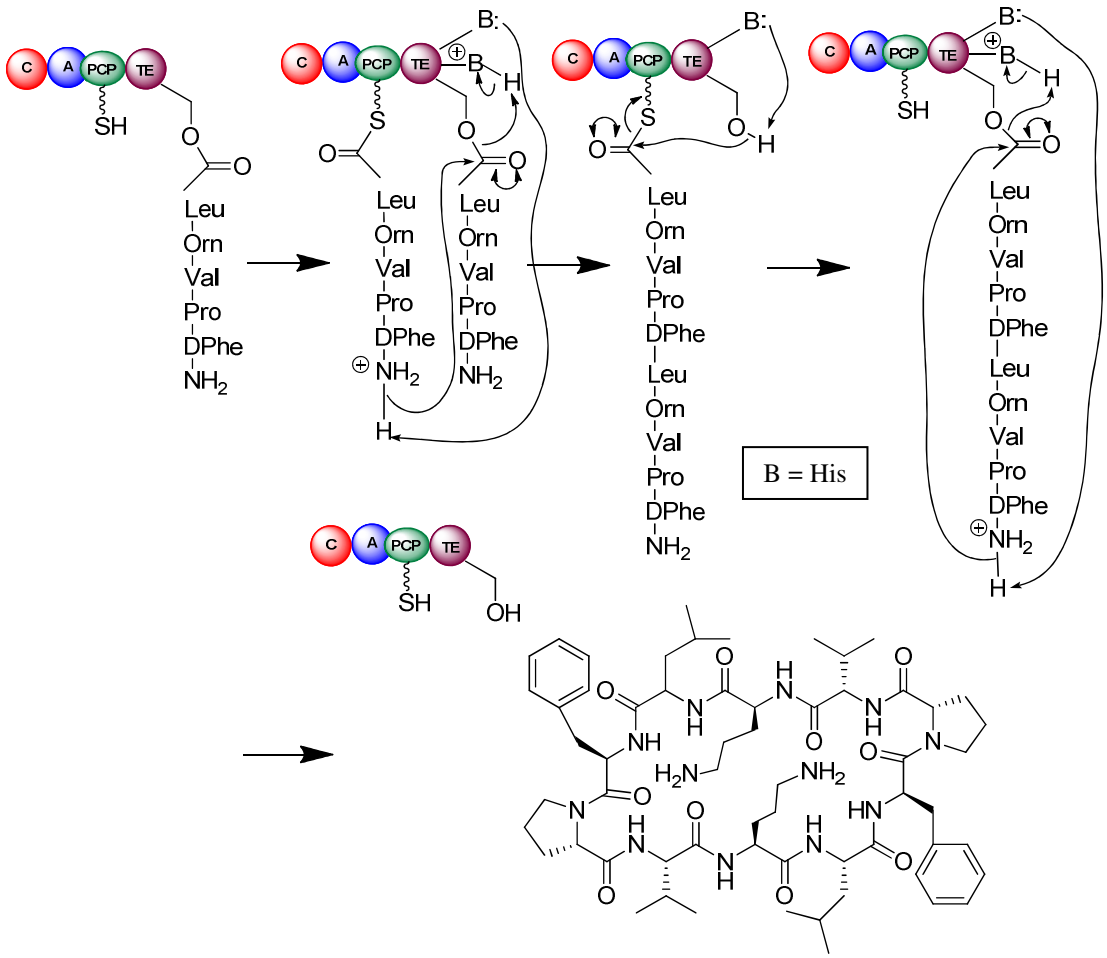


Figure 3-14 Proposed mechanism for cyclooligomerisation catalysed by the iterative thioesterase domain of the gramicidins NRPS.

Another, relatively rare, type of chain termination reaction is catalysed by reductase (R) domains in some TE-independent NRPSs [227, 229]. For example, in the biosynthesis of nostocyclopeptide A2, an R domain catalyses the reductive cleavage of the PCP thioester to form a reactive linear heptapeptide aldehyde which subsequently undergoes cyclisation with the N-terminal amino group to form an imine (Figure 3-15).

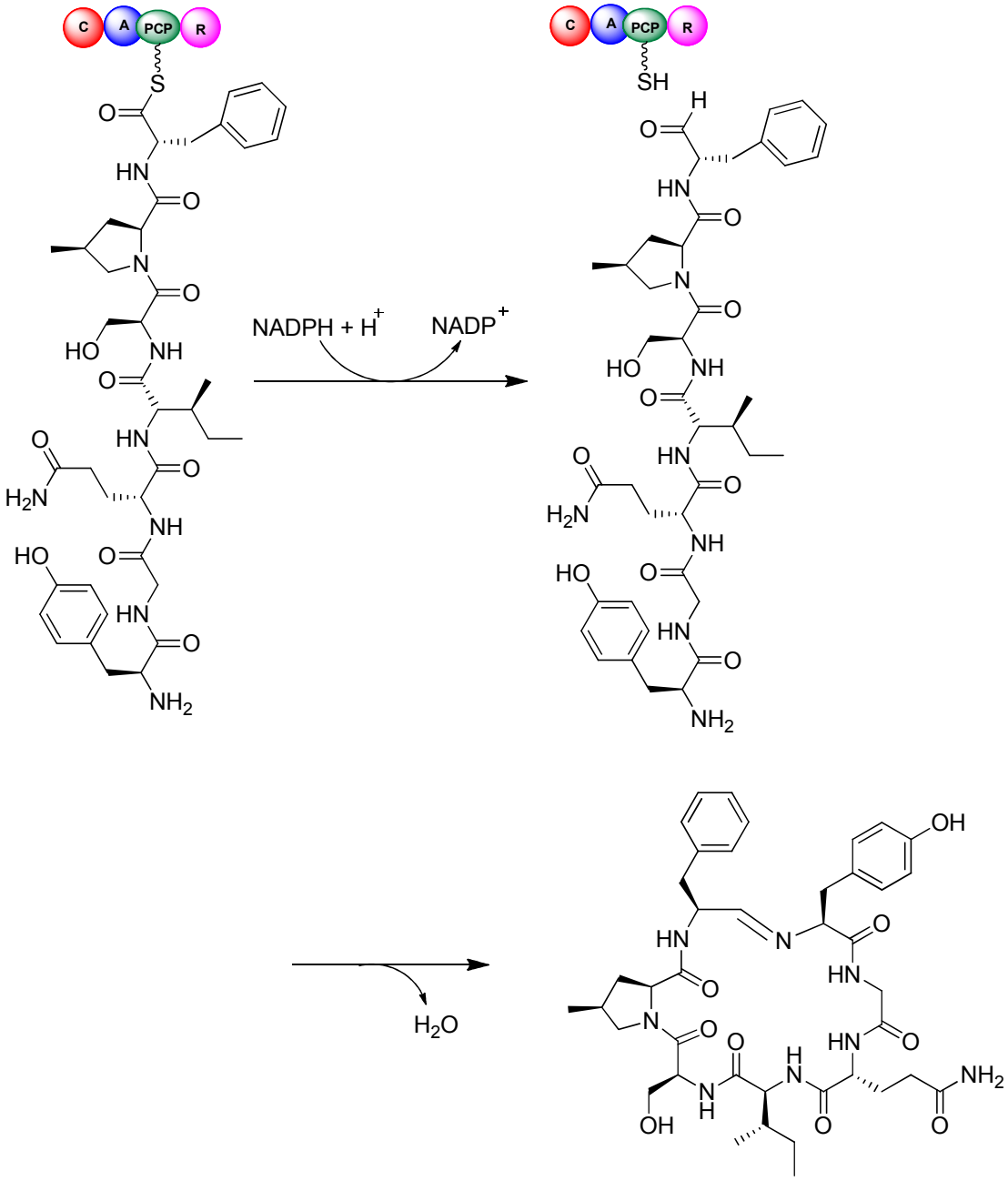


Figure 3-15 Macrocyclisation mechanism catalysed by reductase domain during the biosynthesis of the nostocyclopeptide A2.

C domains have also been reported to be involved in peptide chain release from PCP domains of NRPSs. One of such example is VibH, a condensation domain from the *Vibrio cholera*e vibriobactin synthetase [216, 230]. Unusually, VibH is a free standing condensation domain which catalyses amide bond formation. Another unusual feature of VibH is that one of its substrates is a soluble amine, rather than a PCP-bound amino acid (Figure 3-16).

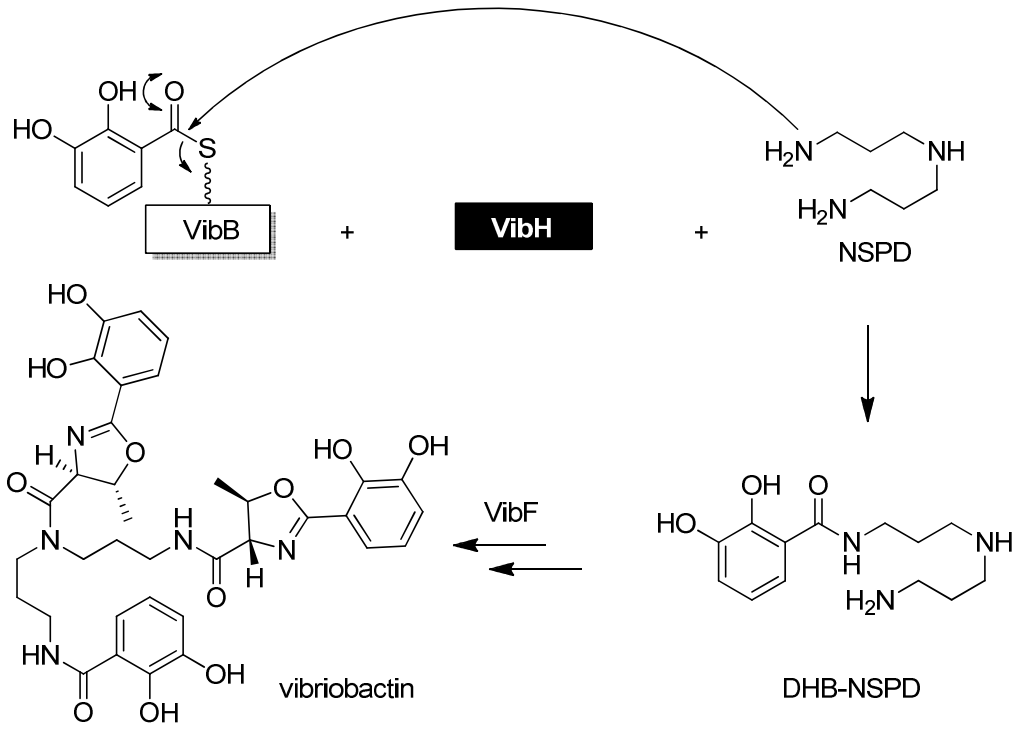


Figure 3-16 VibH, a free standing condensation domain from the vibriobactin synthetase, which catalyses the transfer of 2,3-dihydroxybenzoate (DHB) from the phosphopantetheine of carrier protein VibB to a primary amine of norspermidine (NSPD). The final product, vibriobactin, results from sequential attachment of two 2-(2,3-dihydroxyphenyl)-5-methyloxazoline-4-carbonyl groups catalysed by the six-domain NRPS VibF

The structures of a few NRPS TEs have been reported in the literature including those involved in the biosynthesis of surfactin (Srf) [231], fengycin (Fen) [232] and enterobactin [233]. Structural data has revealed that TE domains belong to the α/β-hydrolase superfamily of enzymes including lipases, proteases and esterases. Typically, NRPS TE domains have a conserved catalytic triad of Asp, His and Ser. In all the TE structures there is a distinctive region called the 'lid region'. The lid region can significantly differ between single TE domains. In the case of the Srf TE (Figure 3-17), this region is extended over three α-helices and can adopt two distinct conformations: open (when the lid is folded back, exposing the active site) and closed (when the lid is covering the active site). On the other hand, the lid region of the Fen TE is very short and so the catalytic triad is exposed all the time. It is suggested that the lid region in this case, plays an indirect role in stereospecific substrate recognition [227]. A PCP-TE didomain structure from the enterobactin synthetase EntF has also been reported [233]. This shows a similar structure, with the catalytic triad and the lid region, to the TEs involved in surfactin and fengycin biosynthesis.

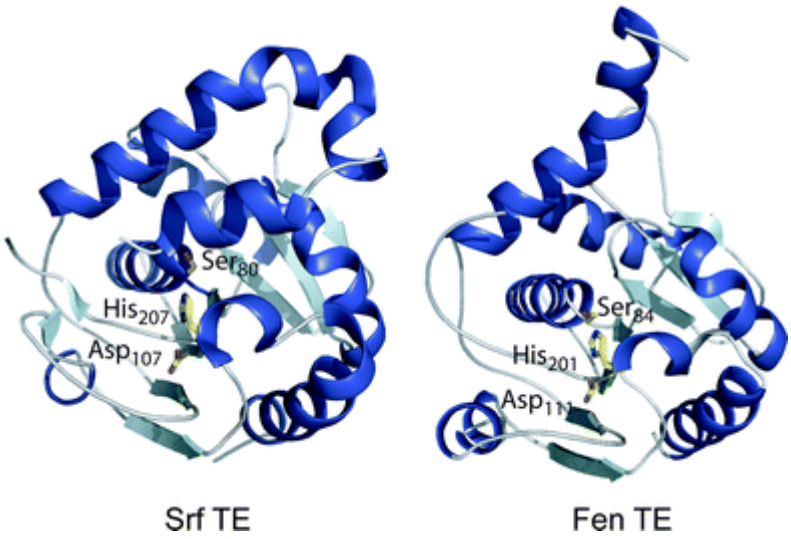


Figure 3-17 Crystal structures of the TE domains from the surfactin synthetase (Srf TE) and the fengycin synthetase (Fen TE). The conserved catalytic triad is highlighted in yellow [231, 232]

3.3.6 The co-linearity rule for NRPSs

The number of modules in an NRPS is usually equal to the number of residues incorporated into the peptide product. This rule of thumb is called the "co-linearity rule", but recently researchers have discovered that some NRPSs do not obey this rule. One exception in *Mycobacterium smegmatis* is the NPRS responsible for assembly of the pentapeptide siderophore exochelin MS (Figure 3-18) which contains six modules. Another example is bleomycin (Figure 3-18). Although this peptide contains 10 amino acid-derived residues, the NRPS responsible for its assembly contains 11 modules [234].

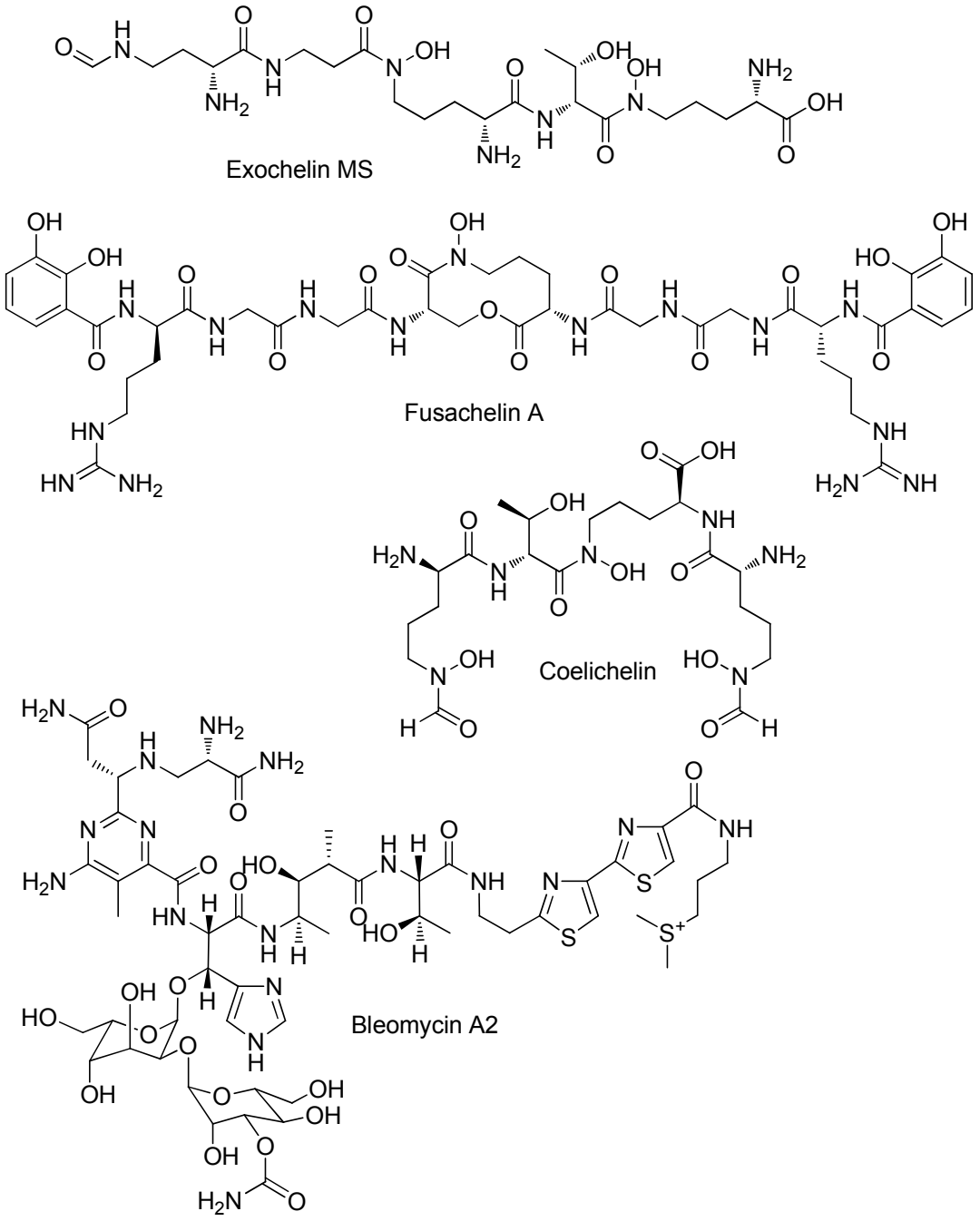


Figure 3-18 Structures of exochelin MS, fusachelin A, coelichelin and bleomycin A2.

The tetrapeptide siderophore coelichelin is another exception to the co-linearity rule. The NRPS responsible for its assembly contains only 3 modules [83]. Yet another exception is fuscachelin A, the probable product of a cryptic gene cluster identified in the genome sequence of the moderate thermophile *Thermobifida fusca* [210]. The NRPS proposed to catalyse fuscachelin biosynthesis contains six modules, yet fusachelin is assembled from closely-related hexapeptide and tetrapeptide chains. It is proposed

that the fifth module of the NRPS is skipped in the ligation of the tetrapeptide and hexapeptide chains.

3.4 The Biosynthesis of coelichelin

Coelichelin is a nonribosomal peptide siderophore produced by *S. coelicolor* M145 that has high structural similarity to foroxymithine (section 3-2). Siderophores are small molecules produced by virtually all microorganisms that have a high affinity for ferric iron. Their role is to scavenge iron from the environment and transport it into the cell through specific receptors. The biosynthesis of siderophores is typically induced by intracellular iron deficiency. More than 500 siderophores have been structurally characterised and interest in the biosynthesis of siderophores is growing because they often function as virulence factors in pathogens. As a consequence, enzymes involved in siderophore biosynthesis are potential antimicrobial drug development targets [234].

Analysis of the genome sequence of *S. coelicolor* M145, published in 2002 [74], revealed numerous gene clusters encoding NRPSs that were not associated with known metabolites. Sequence analyses of one such gene cluster, the *cch* cluster, showed that it encodes 11 proteins, including a trimodular NRPS (CchH) (Figure 3-19) [74]. The substrates for modules 1, 2 and 3 were predicted using structure-based sequence comparisons [83, 86, 235] to be L-δ-N-formyl-δ-N-hydroxyornithine (L-hfOrn), L-threonine (L-Thr) and L-δ-N-hydroxyornithine (L-hOrn), respectively.

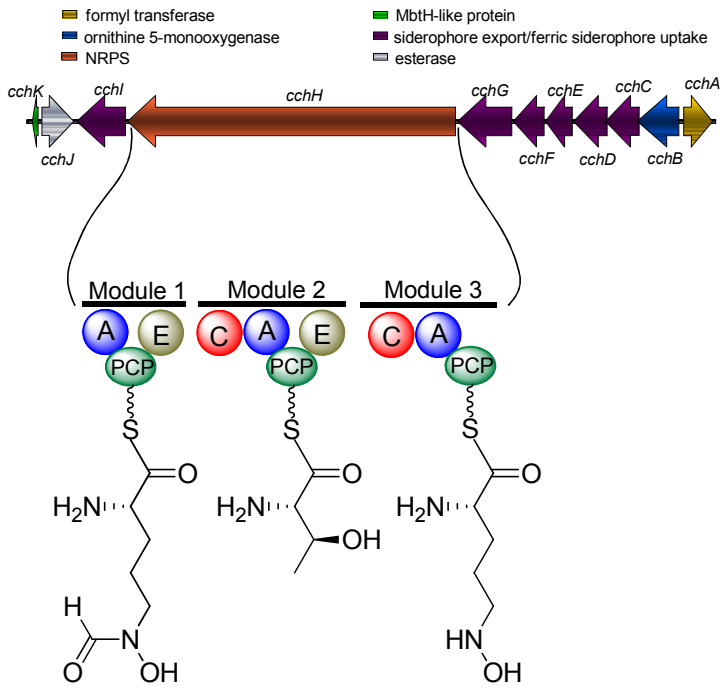


Figure 3-19 Organisation of the coelichelin biosynthetic gene cluster and the module and domain organisation of the NRPS encoded by *cch*H.

The metabolic product of this cryptic biosynthetic gene cluster was identified using a combination of gene knockout and comparative metabolic profiling methodology. The structure of the metabolite, called coelichelin, was elucidated using a combination of high resolution and tandem mass spectrometric analysis of desferri-coelichelin and 1- and 2-D high-field NMR analysis (^{1}H, COSY, TOCSY, HMBC, and ROESY experiments) of the gallium-coelichelin complex. Based on sequence analyses and the finding that *cchH* and *cchJ* are both required for metabolite production, a model for the assembly of the tetrapeptide, coelichelin, by a trimodular NRPS was proposed (Figure 3-20). This model involves sequential condensation of the thioesters of the three substrates, L-hfOrn, L-Thr and L-hOrn to form a tripeptide tethered to the PCP domain of module 3. A second molecule of L-hfOrn is activated by module 1, and condensed with the tripeptidyl thioester using the C domain of either module 2 or 3, skipping the A and T domains of module 2 and the C domain of either module 2 or module 3 in this second iteration of chain assembly [83, 234]. The putative α,β-hydrolase CchJ was proposed to act as a stand alone thioesterase that catalyses hydrolytic release of the tetrapeptide.

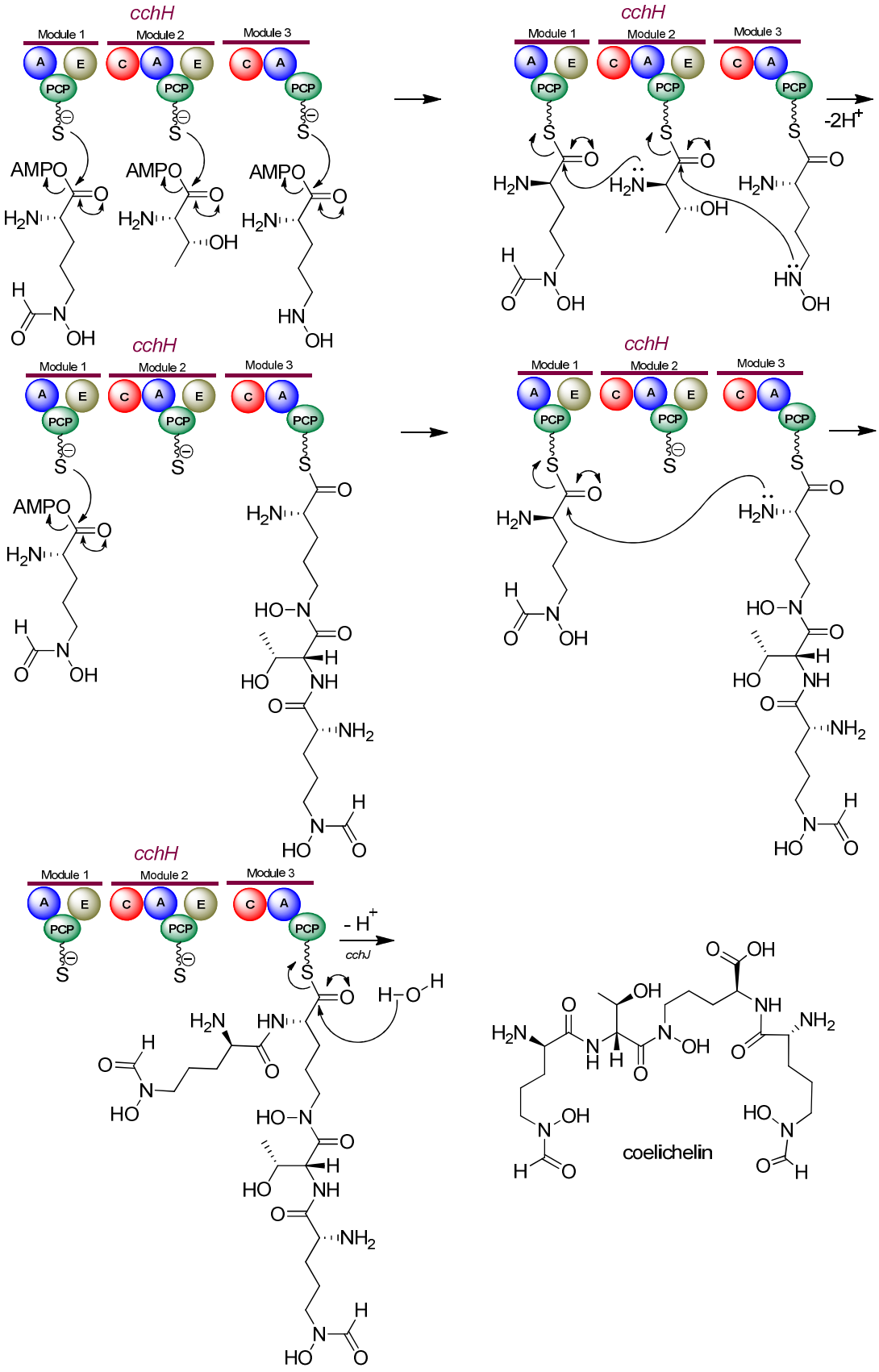


Figure 3-20 Model for coelichelin biosynthesis in *S. coelicolor* involving iterative module use and module skipping by the NRPS CchH.

Two other enzymes CchA and CchB encoded by genes within the coelichelin biosynthetic gene cluster were found to have sequence similarity to formyl transferase and FAD-dependent monoxygenases, respectively. These enzymes were proposed to catalyse hydroxylation and formylation of L-Orn to produce L-hOrn and L-fhOrn, the two non-proteinogenic amino acid residues incorporated into coelichelin. Marahiel and co-workers have experimentally confirmed that CchB is a NADPH and FAD-dependent enzyme that catalyses δ-N-hydroxylation of L-Orn. The role of CchA in δ-N-formylation of δ-N-hydroxyornithine has not yet been confirmed experimentally [234, 236]. However, CchA shows sequence similarity to other formyl transferases known to be involved in nonribosomal peptide biosynthesis [234, 236].

Analysis of the sequence of the coelichelin NRPS showed that modules 1 and 2 contain epimerisation (E) domains. As discussed above, such domains catalyse epimerisation of the α-carbon in aminoacyl thioesters, suggesting that the α-carbons of the substrates of modules 1 and 2 are converted to the D-configuration during peptide chain assembly. The relative stereochemistry of the four α-carbons in coelichelin was determined using molecular modelling of Ga-coelichelin in combination with inter-residue distances and dihedral angles deduced from NOESY and ^{1}H NMR data, respectively, for the metal-siderophore complex. The relative configuration of the α and β-carbons of the threonine residue was determined by acid hydrolysis of Ga-coelichelin followed by derivatisation of the liberated threonine and comparison on a chiral gas chromatography column with authentic standards. On the basis of these studies, the relative stereochemistry of coelichelin was proposed to be D-hfOrn-D-a*llo*-Thr-L-hOrn-D-hfOrn. Although the absolute stereochemistry of coelichelin remains to be experimentally-determined, biosynthetic considerations suggest that it is D, D, L, D.

3.5 Stereochemical analysis of peptides

The interest in the field of stereochemistry started with the discovery of the optical isomerism of tartaric acid by Pasteur in 1848. The physiological environment within a living organism is chiral and as a consequence, the biological activity of enantiomeric forms of molecules can differ. There are 20 ubiquitous amino acids in living

organisms that are the building blocks of proteins and all (with the exception of glycine) have a stereocentre at the α-position. As a consequence, these amino acids can exist as pairs of enantiomers [237].

Discrimination of the enantiomers of amino acids and other chiral molecules has long posed an analytical challenge. Two different approaches have evolved for the separation of enantiomers: the indirect method and the direct method. In the indirect method, the pair of enantiomers is first derivatised with a homochiral reagent to convert them to diastereomers, which are then separated by chromatography on an achiral column. Diastereomers have different physical properties. This is the reason they can be separated using conventional chromatography methods [237, 238]. The direct approach does not involve any chemical derivatisation prior to separation. Methods used in this approach include ligand-exchange chromatography, separation on a homochiral stationary phase [239] and the addition of homochiral compounds to the mobile phase [240].

Each strategy has its advantages and disadvantages, but in general the indirect method, which involves pre-column derivatisation, is less costly, simpler to perform, more easily optimised and gives better resolution than the direct approach.

3.5.1 Marfey's Method for the separation of enantiomeric amino acids

Marfey's reagent (1-fluoro-2,4-dinitrophenyl-5-L-alanine amide) has been widely used for the separation of amino acid enantiomers [238]. The synthesis and application of the Marfey's reagent (Figure 3-21) was first reported in 1984. Enantiomeric mixtures of five amino acids, Ala, Asp, Glu, Met and Phe were separated on a reverse phase HPLC column after derivatisation with the reagent and, in all cases, the diastereomers derived from the L-amino acids were reported to elute before the diastereomers derived from the D-amino acids [241].

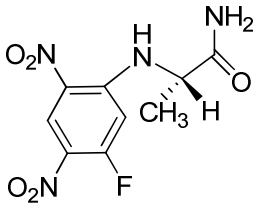


Figure 3-21 Structure of Marfey's reagent (1-fluoro-2,4-dinitrophenyl-5-L-alanine amide, FDAA).

Marfey's reagent has a stereocentre in its alanine-derived moiety. It reacts via nucleophilic aromatic substitution of the fluoro group with the α-amino group of an amino acid, yielding diastereomers (Figure 3-22). Marfey's derivatives have been separated by thin layer chromatography (TLC). Ruterbories and Nurok resolved 22 DL-amino acids on C18 silica layers, but it was not possible to resolve all of these amino acids on one TLC plate because of overlapping R_f values. [242]. Reverse phase-TLC plates have also been used to separate Marfey's derivatives of DL-Glu, DL-Asp, DL-Thr, DL-Pro and DL-Ala [243]. However Marfey's derivatives are most conveniently and easily separated on a C-18 HPLC column.

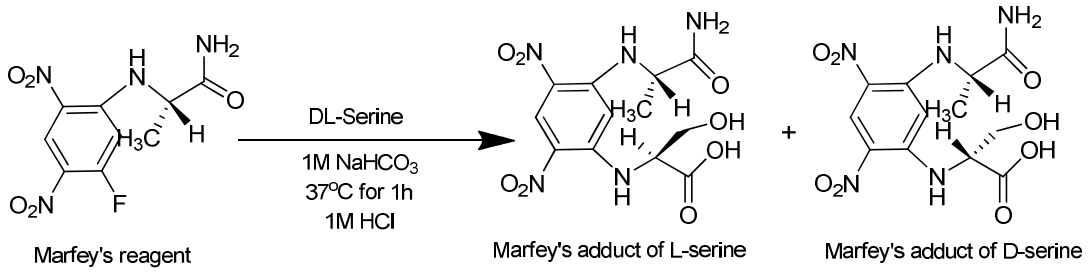


Figure 3-22 Diastereomeric products resulting from the reaction of DL-Ser with Marfey's reagent.

One disadvantage of Marfey's method is that when diamino acids such as Orn and Lys are derivatised, they can form both mono- and di-substituted Marfey's derivatives, which can complicate subsequent analyses. However, the use of a slight excess of Marfey's reagent can minimise di-substitution [244]. Racemisation during any reaction involving homochiral starting materials is a concern and has the potential to seriously affect the outcome of stereochemical analyses. Racemisation of amino acids

can occur during the acid hydrolysis of peptides [245] but no racemisation was detected when Marfey's reagent was reacted with L-Ala, L-Ser and L-Pro [246, 247].

Aims and objectives

The planar structure of foroxymithine closely resembles that of coelichelin (section 0) suggesting both molecules might share similar biosynthetic pathway. The relative stereochemistry of the four α-carbons (N-terminal→C-terminal) of coelichelin D-D-L-D, has been reported whereas the stereochemistry of the α-carbons of each of the four amino acids residues within foroxymithine has been reported to be L. However, in the case of foroxymithine, no experimental data in support of this stereochemical assignment have been published. Moreover the published ^{1}H NMR data for foroxymithine isolated from *Streptomyces nitrosporeus* differs significantly from that reported for synthetic foroxymithine. The aim of this project was to investigate the structure, stereochemistry and biosynthesis of foroxymithine isolated from *Streptomyces narbonensis*, which has not been reported as a producer of this natural product previously. The main objectives of the project were to:

1) Purify foroxymithine from *S. narbonensis*
2) Elucidate the structure of foroxymithine using 1- and 2-D NMR spectroscopy, together with high resolution and tandem mass spectrometry..
3) Investigate the relative stereochemistry of foroxymithine using NMR spectroscopic analysis and computer modelling of its gallium complex.
4) Determine the absolute stereochemistry of each of the amino acid residues within foroxymithine by chromatographic comparison of the Marfey's derivatives of the amino acids liberated by acid-catalysed hydrolysis of the natural product with Marfey's derivatives of commercially available D- and L-serine and chemically synthesised D- and L-δ-N-hydroxyornithine.
5) Use degenerate PCR primers to amplify fragments of putative foroxymithine biosynthetic genes from genomic DNA of *S. narbonensis*.
6) Construct a genome fosmid library of *S. narbonensis* and screen it for clones containing the putative foroxymithine biosynthetic gene fragments.

Part II: Structural and Biosynthetic Investigations of Foroxymithine

CHAPTER 4 - RESULTS & DISCUSSION

4.1 Purification of foroxymithine from *Streptomyces narbonensis* and ESI-MS analysis

Streptomyces narbonensis was grown in an iron deficient medium to ensure high levels of metabolite production. After incubation for 4 days at 30°C the culture was filtered, and the supernatant was passed through a HP-20 resin column and eluted with 20% methanol. The eluate was lyophilised and the residue was redissolved in water and $FeCl_3$ was added to convert foroxymithine to its ferric complex. The ferric complex of foroxymithine was made because it is much more stable than the free form and also due to high UV absorbance of the Fe-hydroxamate complex chromophore, facilitating its separation on HPLC.

A semi-preparative C-18 HPLC column was used to purify ferri-foroxymithine. The column was eluted using a gradient of the mobile phases H_2O (containing 0.1% TFA) and acetonitrile (containing 0.1% TFA). The elution of ferric hydroxamates was monitored by observing the change in absorbance at 435 nm. Fractions containing compounds at this wavelength were collected manually (Figure 4-1). The retention time of ferri-foroxymithine was around 8-10 minutes which is comparable to the retention time of 7-9 minutes reported for purification of ferri-coelichelin under similar conditions [248]. The presence of ferri-foroxymithine in the collected fractions was confirmed by positive ion mode ESI-MS. An ion with *m/z* = 629 indicated the presence of protonated ferri-foroxymithine. Ions with *m/z* = 627, 630 and 631 flanking the *m/z* 629 ion were consistent with an iron-containing species. Mono-sodiated ferri-foroxymithine was also observed giving rise to an ion with *m/z* = 651 (Figure 4-2).

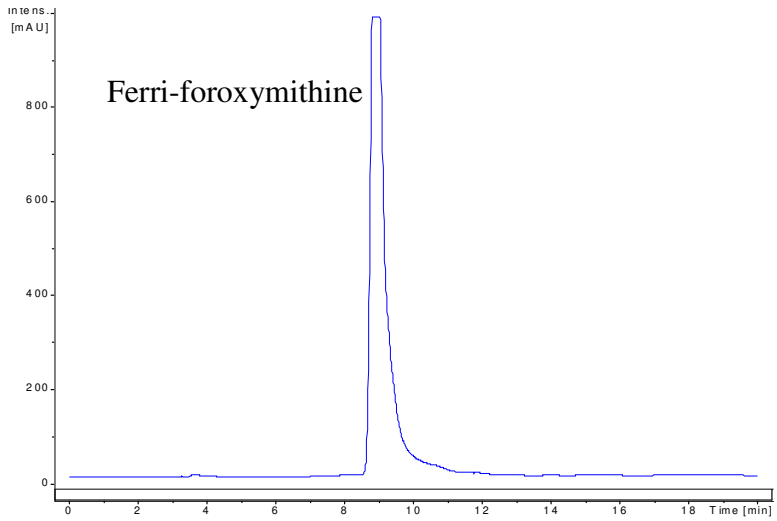


Figure 4-1 Chromatogram for purification of ferri-foroxymithine from concentrated culture supernatant of *S. narbonensis* by HPLC monitoring absorbance at 435 nm

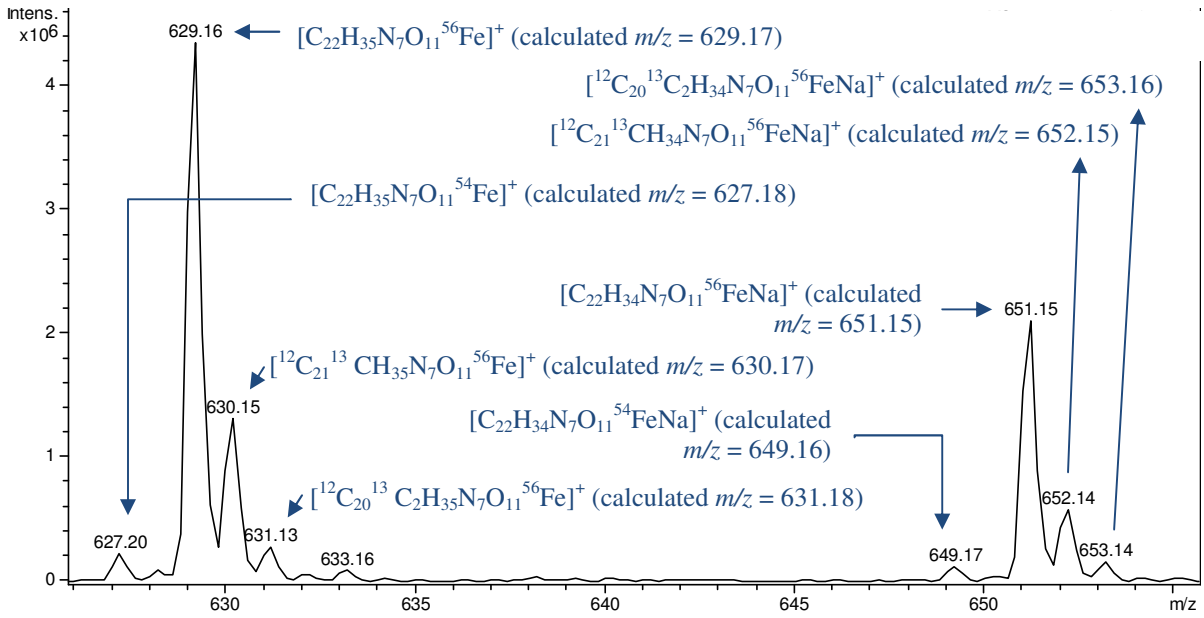


Figure 4-2 Positive ion ESI-mass spectrum of purified ferri-foroxymithine

In the next step iron was removed from the ferri-foroxymithine complex by the method of Lautru *et al* [83]. All glassware used was washed with concentrated nitric acid followed by thorough washing in deionised water to remove ferric ions from the surfaces of the glassware. 8-hydroxyquinoline was used as a chelating agent to remove iron from the ferri-foroxymithine complex [83, 249].

A large excess of 8-hydroxyquinoline was dissolved in methanol and added to a dilute aqueous solution of ferri-foroxymithine. A dark-green Fe-(8-hydroxyquinoline)$_3$ complex was formed, which was extracted using dichloromethane. The colourless aqueous phase was separated and lyophilized. The presence of desferri-foroxymithine in the residue was confirmed by the presence of an *m/z* = 576 ion in a solution of the residue analysed by positive ion mode on ESI-MS (Figure 4-3).

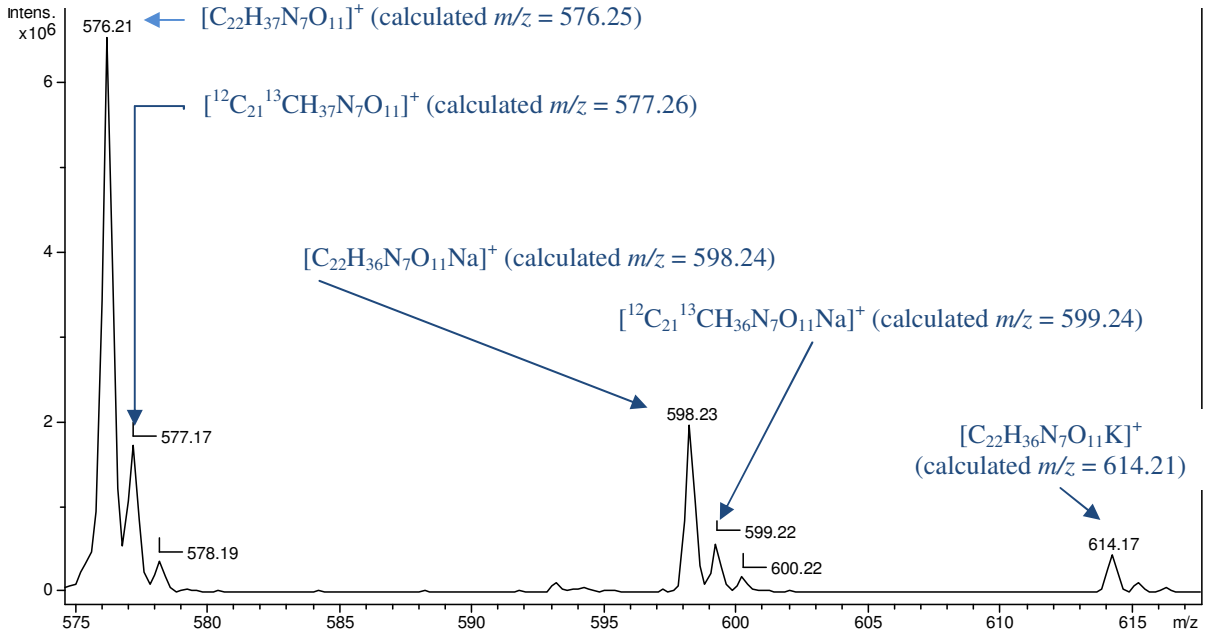


Figure 4-3 ESI-mass spectrum of desferri-foroxymithine after purification by HPLC

In the next step, the Ga-complex of foroxymithine was made by mixing a solution of gallium sulphate in sulphuric acid with an aqueous solution of lyophilised desferri-foroxymithine for 30 minutes at room temperature. The solution was neutralised with NaOH and lyophilised after snap freezing in liquid nitrogen. The residue was dissolved in MilliQ water and purified by HPLC on a C18 column using isocratic 10 mM ammonium carbonate buffer adjusted to pH 7, as the mobile phase and monitoring absorbance at 210 nm. (Figure 4-4).

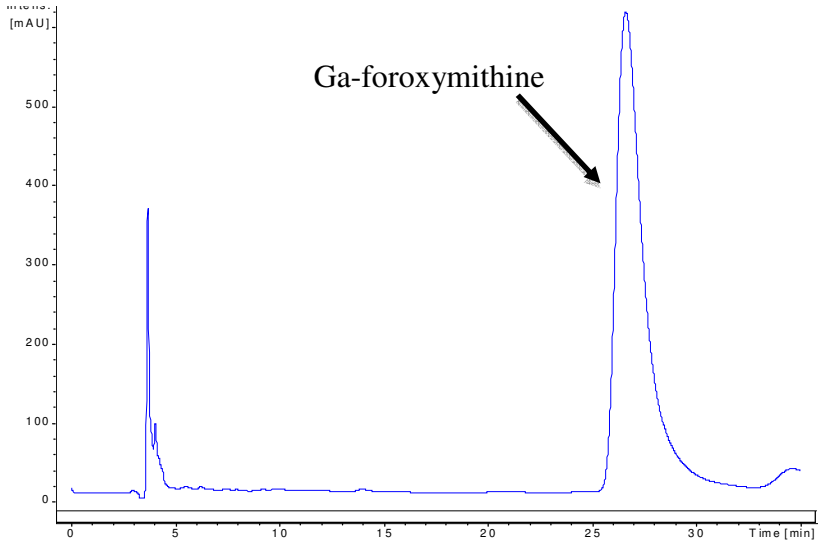


Figure 4-4 Chromatogram from HPLC purification of gallium-foroxymithine complex monitoring absorbance at 210 nm.

A fraction corresponding to the peak eluting at approximately 25 minutes was collected and analysed by ESI-TOF-MS. The presence of gallium-foroxymithine was confirmed by the observation of ions with *m/z* = 642 ([Ga-foroxymithine + H]$^+$) and *m/z* = 664 ([Ga-foroxymithine + Na]$^+$) (Figure 4-5).

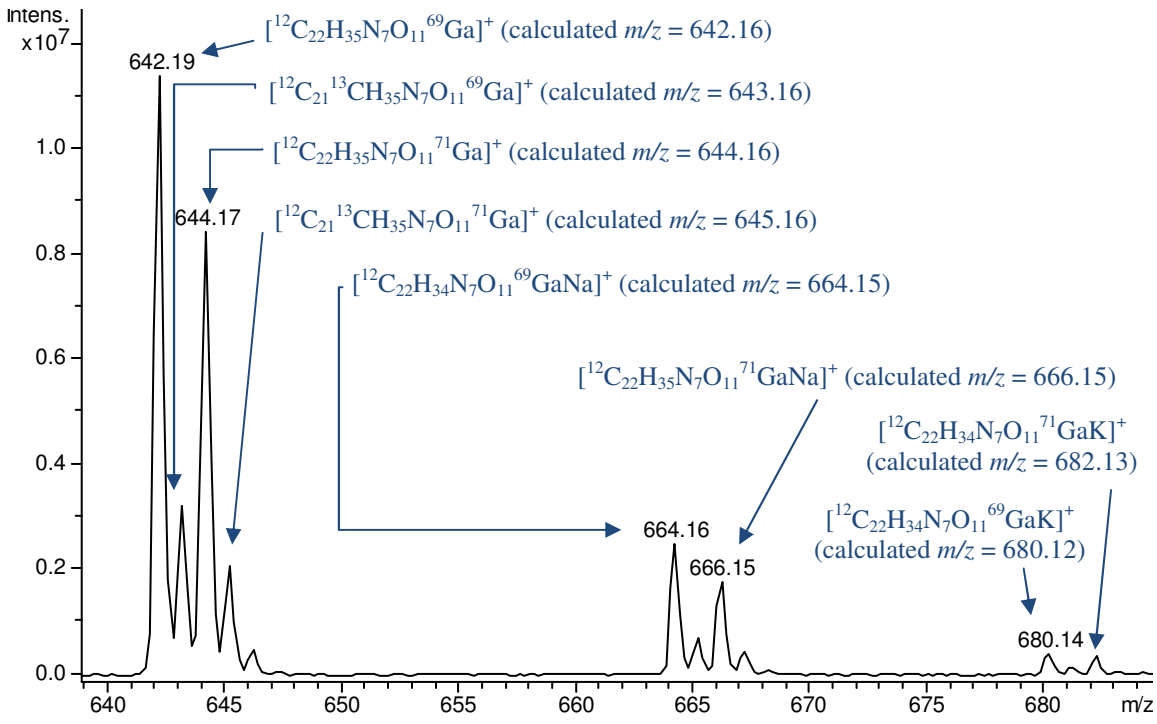


Figure 4-5 ESI-mass spectrum of gallium-foroxymithine after purification by HPLC.

High resolution positive ion mode ESI-TOF-MS was carried out to confirm the molecular formula of the Ga-foroxymithine complex (Figure 4-6).

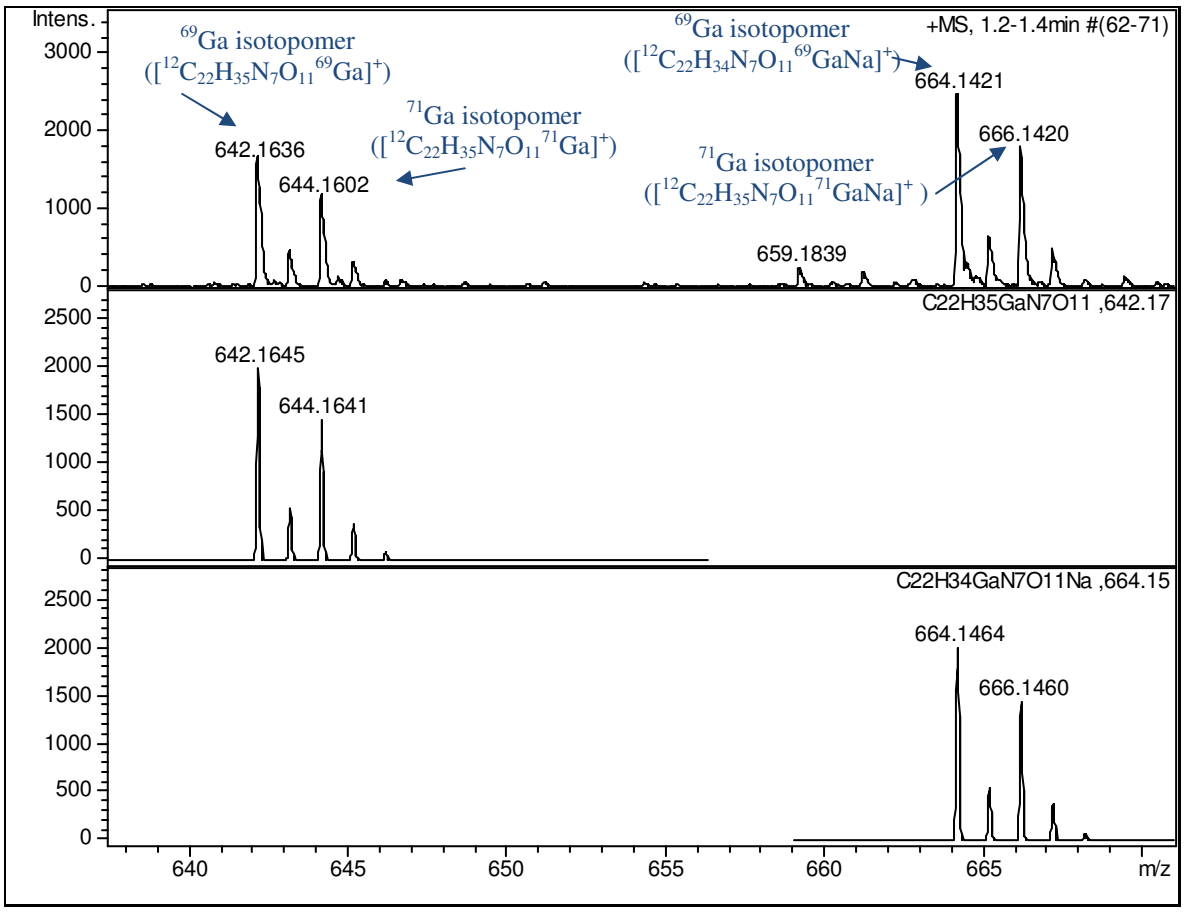


Figure 4-6 ESI-TOF mass spectrum of Ga-foroxymithine (top); simulated mass spectrum for the [M+H]+ ion of Ga-foroxymithine (middle); simulated mass spectrum for the [M+Na]+ ion of Ga-foroxymithine (bottom)

The gallium-foroxymithine complex was subjected to ESI-MS/MS. Fragmentation of the molecular ion (*m/z* = 642) produced an ion with *m/z* = 624 by loss of water. Further fragmentation led to loss of another water molecule to give an *m/z* = 606 ion which in turn gave an *m/z* = 588 ion by the loss of a further water molecule. Major ions produced by fragmentation of the *m/z* 588 ion had *m/z* = 542, 472, 406 and 337 (Figure 4-7 & Table 4.1).

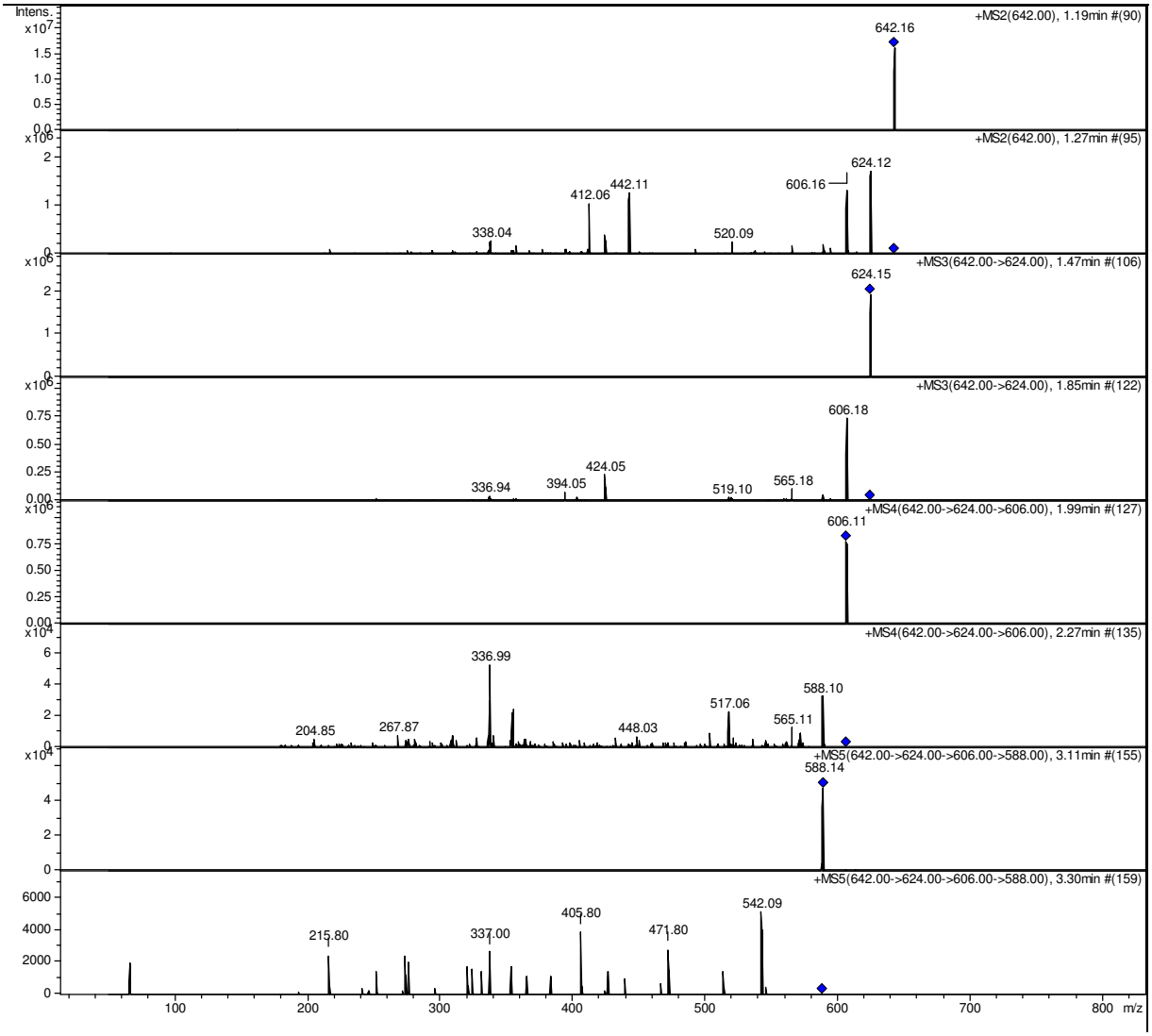


Figure 4-7 ESI-MS/MS spectra showing fragmentation pattern for the Ga-foroxymithine [M+H]+ ion (*m/z* = 642)

Table 4.1 Possible fragment ions derived from the parent ion of (*m/z* = 642) for Ga-foroxymithine (Ga-fm)

m/z	Possible ion	Molecular formula
642	Ga-fm + H^+	$C_{22}H_{34}GaN_7O_{11}^+$
624	Ga-fm + H^+ - H_2O	$C_{22}H_{32}GaN_7O_{10}^+$
606	Ga-fm + H^+ - $2H_2O$	$C_{22}H_{30}GaN_7O_9^+$
588	Ga-fm + H^+ - $3H_2O$	$C_{22}H_{28}GaN_7O_8^+$
542	Ga-fm + H^+ - $C_4H_2O_3$	$C_{18}H_{30}GaN_7O_8^+$
472	Ga-fm + H^+ - $C_4H_3GaO_3$	$C_{18}H_{31}N_7O_8^+$
406	Ga-fm + H^+ - $C_8H_{16}N_2O_6$	$C_{14}H_{19}GaN_5O_5^+$
337	Ga-fm + H^+ - $C_{11}H_{19}N_3O_7$	$C_{11}H_{16}GaN_4O_4^+$

4.2 NMR spectroscopic analysis of the Ga-foroxymithine complex

The planar structure of foroxymithine was confirmed using 1D and 2D NMR experiments. Comparison of the ^{1}H NMR spectrum of foroxymithine isolated from *S. narbonensis* with the ^{1}H NMR data reported in the literature for synthetic foroxymithine [176] were in good agreement. However, the ^{1}H NMR data were not in full agreement with the ^{1}H NMR data reported for foroxymithine isolated from *S. nitrosporeus* [55]. Comparison of the ^{1}H NMR chemical shift assignments for foroxymithine from the three different sources is given in Table 4.2. As can be seen from this comparison, it is not possible to conclude beyond reasonable doubt that the three compounds are identical. Thus we sought to independently confirm the structure of foroxymithine isolated from *S. narbonensis*.

Table 4.2 Comparison of the ^{1}H NMR chemical shift assignments for foroxymithine from three different sources

^{1}H location	Foroxymithine from *S. narbonensis* ($(CD_3)_2SO$ as solvent)	Synthetic foroxymithine[176] (D_2O as solvent)	Foroxymithine from *S. nitrosporeus*[55] (D_2O as solvent)
CH_2	12H, m, δ 1.3-1.7	12H, m, δ1.35-1.80	12H, δ 2.0-2.4
CH_3CON	3H, s, δ 1.83	3H, s, δ1.83	3H, δ 2.5
CH_2N	1H, m, δ 3.40-3.45 2H, m, δ 3.80 3H, m, δ 3.35 – 3.40	1H, m, δ 3.00-3.12 2H, m, δ 3.60-3.75 3H, m, δ 3.3.-3.45	δ 4.25-4.44 δ 3.90-4.25
NCHCO	2H, m, δ 3.5 -3.6 1H, m, δ 4.35 1H, m, δ 4.88	2H, m, δ 3.46-3.55 1H, m, δ 4.20-4.30 1H, t, δ 4.88	2H, δ 4.75 1H, δ 4.75-4.90 1H, δ 5.46-5.65
CH_2O	2H, m, δ 3.56	2H, m, δ 3.95-4.05	-
2 x CHO	1H, s, δ 7.97 1H, s, δ 8.23	2H, 3s, δ 7.76, 7.83, 8.11	1.5H, δ 8.42 0.5H, δ 8.78

It was not possible to unambiguously assign all of the signals in the ^{1}H NMR spectrum of foroxymithine isolated from *S. narbonensis*, even with the help of 2D NMR experiments such as COSY, HMBC and HSQC. This is due to the fact that

foroxymithine contains three ornithine-derived residues, giving rise to considerable signal overlap.

To overcome this problem, a complex with Fe^{3+} or Ga^{3+} can be made. The advantage is that the molecular framework becomes more rigid, resulting in better signal dispersion in the ^1H NMR spectrum. A disadvantage of the iron complex is that Fe^{3+} is paramagnetic. This would be expected to result in very broad signals in the proton NMR spectrum. However, Ga^{3+}, which has the same charge, a similar ionic radius in six-coordinate complexes, and similar ligand exchange rates, but is diamagnetic, is an ideal alternative to Fe^{3+}. A similar strategy was employed in the case of coelichelin structural analysis by NMR where it was found that the ^1H NMR spectrum of the Ga^{3+} complex had much greater signal dispersion than the spectrum of the free ligand. Coelichelin also possesses three N5-hydroxyornithine-based residues resulting in considerable signal overlap in the ^1H NMR spectrum of the free ligand. Hence, the Ga^{3+} -foroxymithine complex (Figure 4-8) was made and analysed using 1D and 2D NMR experiments. The ^1H NMR spectrum of this complex showed better dispersion of signals compared to desferri-foroxymithine.

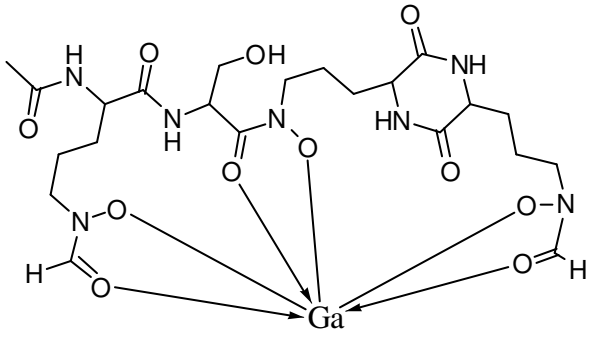


Figure 4-8 Planar structure of Ga-foroxymithine complex

Careful inspection of the ^1H NMR spectrum of Ga-foroxymithine revealed two signals (major and a minor) for several of the protons. (Figure 4-9).

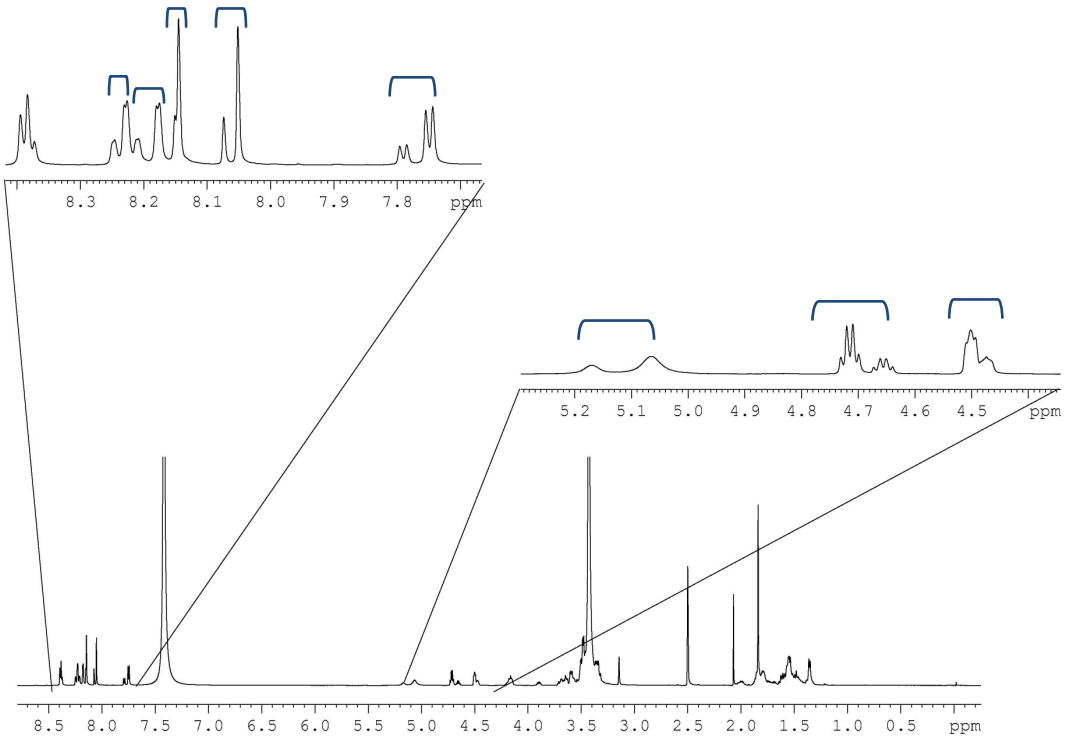


Figure 4-9 ^{1}H NMR spectrum of Ga-foroxymithine recorded at 700 MHz showing signal duplication for several protons

Further analysis of the 1D and 2D NMR spectra of Ga-foroxymithine confirmed the presence of duplicate signals. Both the major and the minor signals correlated with the same carbon atoms in the HMBC spectrum. For example, Figure 4-10 shows an expanded region of the HMBC spectrum containing correlations between the proton signals for Ser-C2 with the carbonyl carbon atoms Ser-C1 and ahfOrn-C1. Such signal doubling has been reported before, for example in the ^{1}H NMR spectrum of the gallium erythrochelin complex [250].

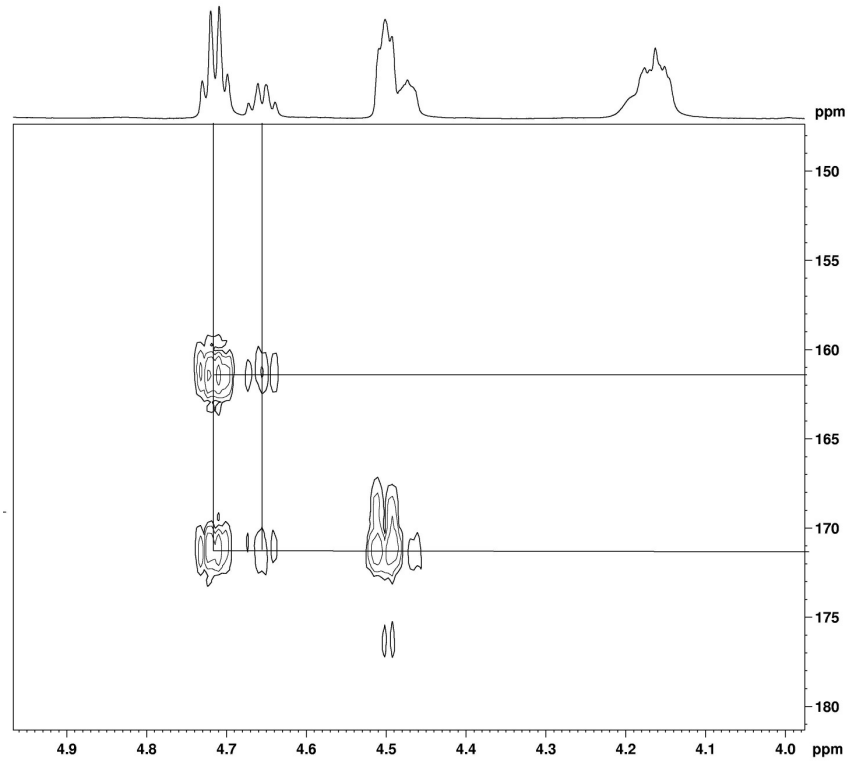
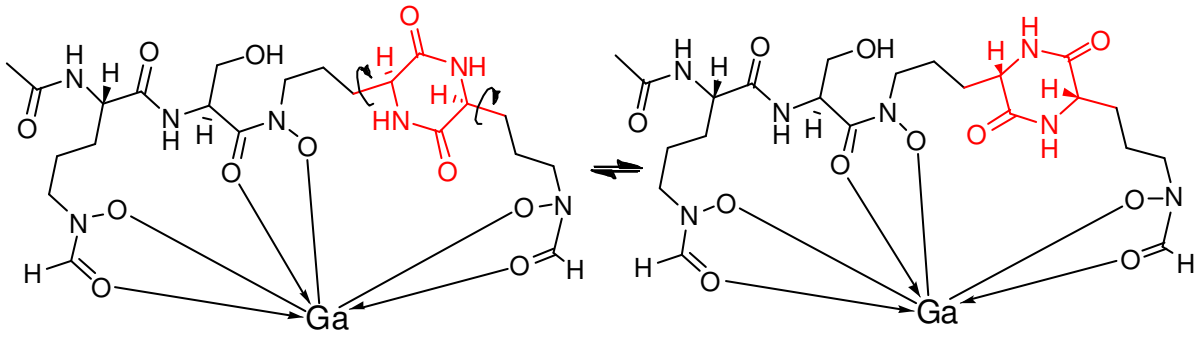


Figure 4-10 Expanded region of the HMBC NMR spectrum of Ga-foroxymithine showing correlations between the duplicate signals due to the Ser-C2 protons with the carbonyl carbon atoms of Ser-C1 and ahfOrn-C1

One possible explanation for the observed signal doubling in the ^{1}H NMR spectrum is the presence of diastereomers resulting from different arrangements of the ligands around the Ga^{3+} (Figure 4-11). Another possible explanation is that there are different conformational isomers present that interconvert slowly at room temperature. The major set of signals in the ^{1}H NMR spectrum may be due to the low energy rotamer and the minor set of signals may result from the higher energy rotamer.

(A)



(B)

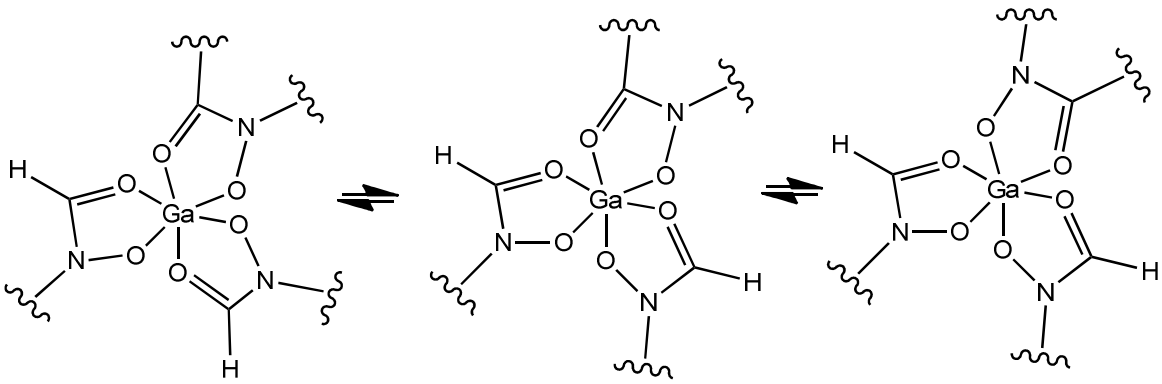


Figure 4-11 Possible isomeric structures of Ga-foroxymithine: A) conformational isomers, B) configurational isomers

4.2.1 LC-MS analysis of Ga-foroxymithine on a homochiral stationary phases

To attempt to resolve the diastereoisomeric forms of Ga-foroxymithine, the complex was analysed by HPLC-MS using different homochiral stationary phases. Two columns were used: a Chiralpak® IA (0.46 cm x 25 cm, Amylose tris-(3,5-dimethylphenylcarbamate immobilised on 5µm silica gel) and a Chiralpak® IC (0.46 cm x 25 cm, cellulose tris-(3,5-dichlororphenylcarbamate) immobilised on 5µm silica gel). Figure 4-12 shows the extracted ion chromatogram at m/z = 642 ($[M + H]^+$ for Ga-foroxymithine) from the analysis using the Chiralpak® 1A column. Figure 4-13 shows the results of an analogous analysis using the Chiralpak IC column. In neither

case was it possible to observe two clearly resolved peaks. Thus it was not possible to separate the two isomeric species observed in the NMR analyses.

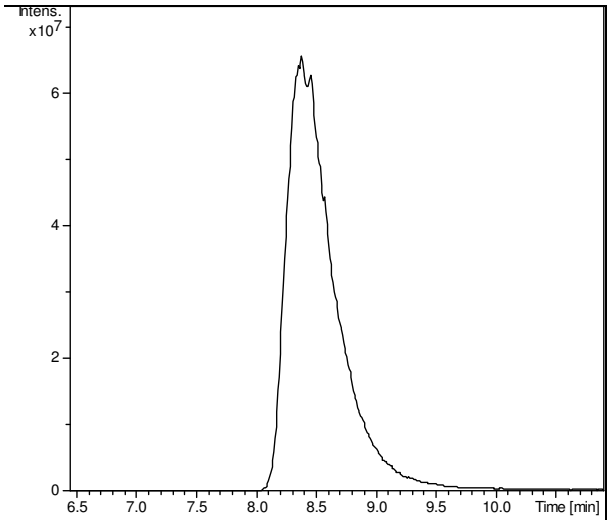


Figure 4-12 Extracted ion chromatogram at *m/z* = 642 from LC-MS analysis of Ga-foroxymithine on the Chiralpak® IA column

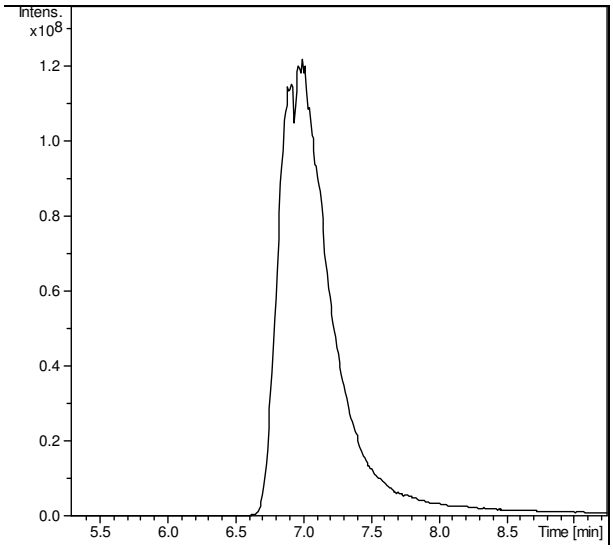


Figure 4-13 Extracted ion chromatogram at *m/z* = 642 from LC-MS analysis of Ga-foroxymithine on the Chiralpak® IC column

4.2.2 ^{71}Ga-NMR Spectroscopic analysis of Ga-foroxymithine

To further investigate the existence of stereoisomeric ligand arrangements around Ga in Ga-foroxymithine, ^{71}Ga-NMR spectroscopy was carried out. If there is more than one stable arrangement of the ligands around the Ga of Ga-foroxymithine it might be possible to see two (or more) signals in the ^{71}Ga NMR spectrum at room temperature.

The ^{71}Ga NMR spectrum of Ga-foroxymithine (Figure 4-14) shows only a single signal, consistent with there being only a single stable configuration of Ga-foroxymithine.

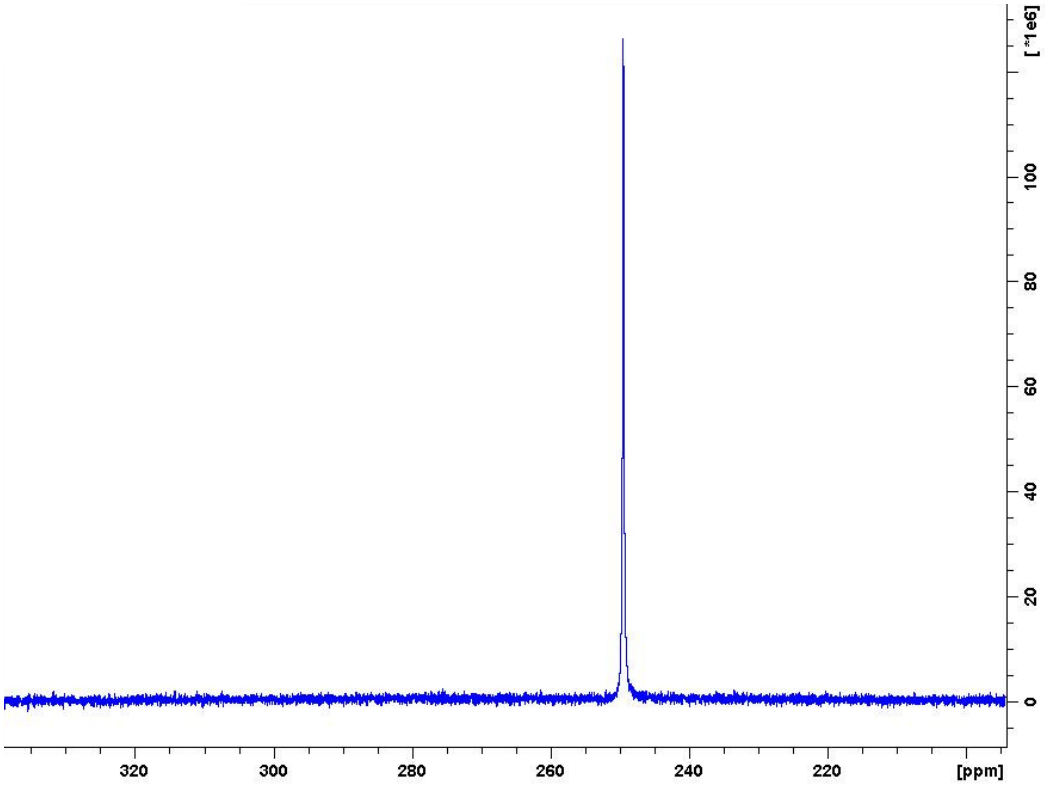


Figure 4-14 ^{71}Ga-NMR spectrum of Ga-foroxymithine

4.2.3 Variable Temperature (VT) – NMR analysis of Ga-foroxymithine

To further investigate the cause of signal doubling variable temperature (VT)-NMR experiments were carried out in which ^{1}H NMR spectra were recorded at different temperatures. As the temperature increased, the rate of interconversion of the isomers increased, such that the duplicated signals eventually coalesced. Figure 4-15 shows the ^{1}H NMR spectra measured at temperatures from 298 to 408K.

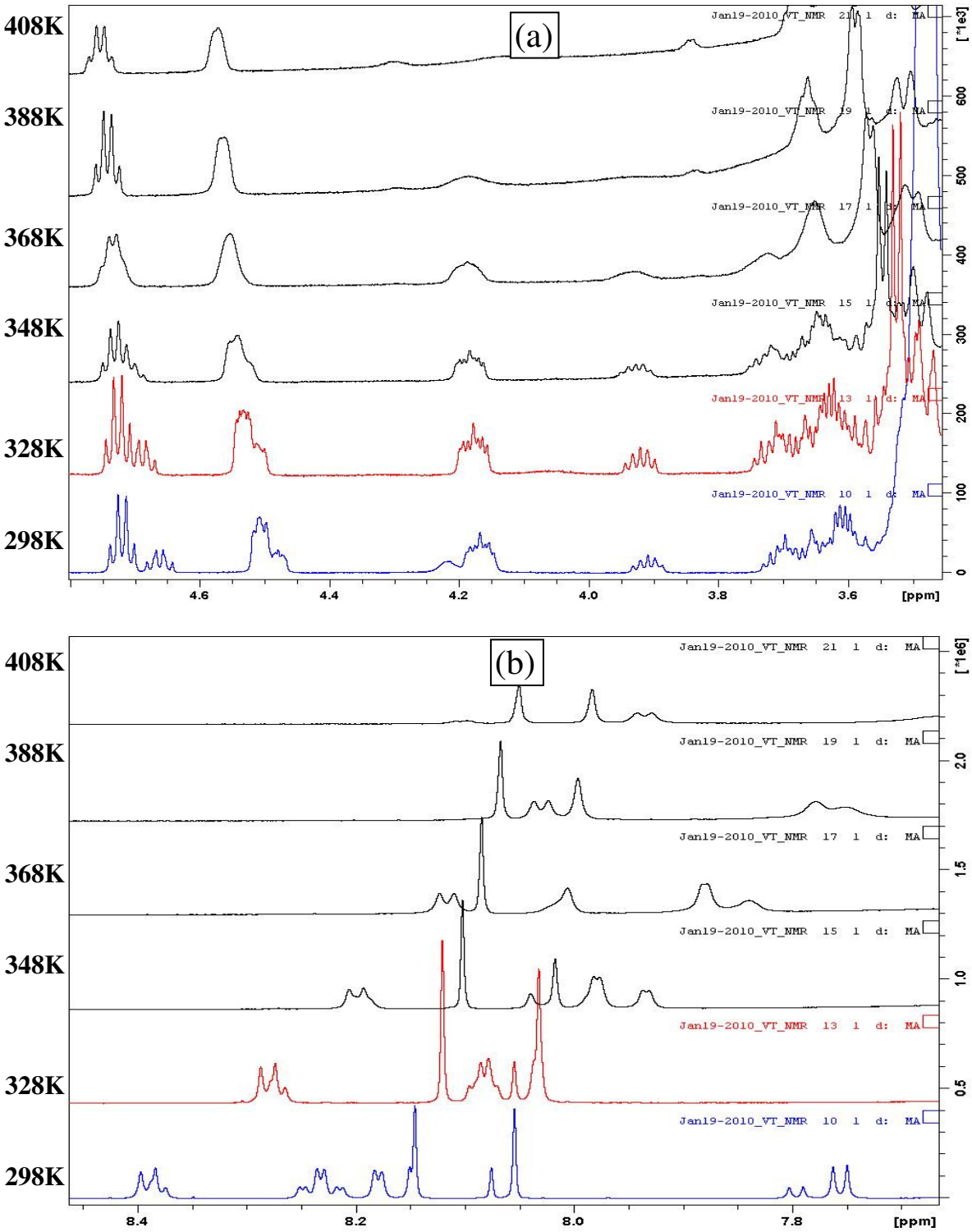


Figure 4-15 ^{1}H NMR spectra of Ga-foroxymithine recorded at 298 – 408K, showing signal coalescence at the a) high field end of the spectrum and b) the low field end of the spectrum.

As can be seen in Figure 4-9 the major and minor signals at ~ 4.5 ppm, which correspond to the α-proton of the N2-acetyl-N5-formyl-N5-hydroxyornithine (ahfOrn) residue gradually coalesce as the temperature increases. At 388K (115°C) the two signals become one. Similarly the major and minor signals at ~ 4.7 ppm coalesced

into one at 388K (115°C). Analogous behaviour was seen at the low field end of the spectrum, where duplicate signals coalesced at around 115°C (388K).

As discussed earlier, one of the possible explanations for duplicate signals in the ^{1}H NMR spectrum of Ga-foroxymithine is the existence of conformational isomers resulting from restricted rotation around single bonds. The results of the VT-NMR spectroscopic analyses are consistent with this explanation. As the temperature increases the rate of rotation about individual bonds increases until it becomes faster than the time scale of the NMR experiment, at which point only a single set of NMR signals is observed.

4.2.4 Structure elucidation of Ga-foroxymithine using 1D and 2D NMR data

Ga-foroxymithine was dissolved in DMSO-d6 and its ^{1}H NMR spectrum was recorded on a Bruker Avance 700 MHz NMR spectrometer equipped with a TCI cryoprobe. For the assignment of the proton signals within each amino acid residue, COSY and TOCSY experiments were carried out. In these experiments proton-proton coupling through bonds can be used to unravel direct (COSY) as well as relayed (TOCSY) connectivities within spin systems (Figure 4-16). Proton-carbon correlation experiments (HMBC, HSQC) were also carried out on the 700MHz NMR instrument. HMBC and HSQC experiments allowed the observation of H, C long-range (2-3 bonds) and short-range (single bond) coupling, respectively (Figure 4-16). The planar structure of foroxymithine was confirmed by careful analysis of NMR data.

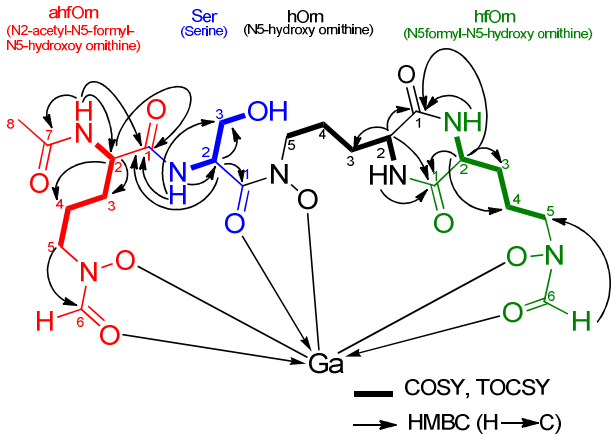


Figure 4-16 Spin systems identified in Ga-foroxymithine by TOCSY and COSY experiments, and correlations observed in the HMBC spectrum

The TOCSY spectrum (Figure 4-17) allowed the identification of four spin systems. The serine spin system was identified by starting from the peak at δ = 8.4 ppm, assigned as the Ser-NH. This signal showed correlations to signals assigned to the Ser C-2, C-3 and hydroxyl group protons. The signals for the corresponding protons within the afhOrn, hfOrn and hOrn residues were assigned similarly.

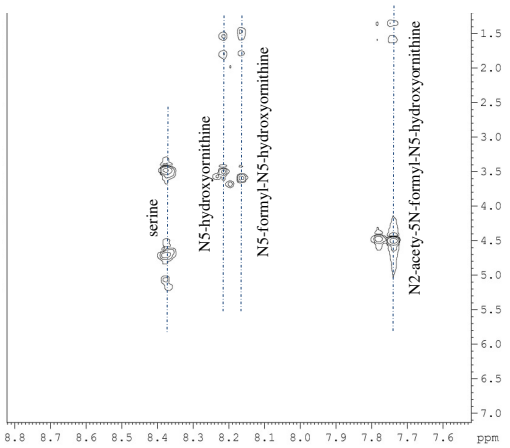


Figure 4-17 Expanded region of the ^{1}H TOCSY spectrum (700 MHz) of Ga-foroxymithine in d6-DMSO. Correlations from the amide protons to each of the other protons within the spin system for each amino acid residue are indicated.

Connectivity among the four amino acid residue was confirmed with the help of signals originating on HMBC NMR spectrum as shown in Figure 3-16. H to C correlation signals from ahfOrn-NH to C-7 and C-2 of ahfOrn and the signal from C-2 proton to C-1 carbonyl carbon of ahfOrn confirmed the location of the acetyl group.

Connection between ahfOrn and Ser residue was confirmed by H to C HMBC signals from protons at ser-NH and Ser C-2 to ahfOrn C-1. The connectivity between hOrn and hfOrn residues was identified from the HMBC signals from hOrn C-2 and hOrn NH to hfOrn C-1. Similarly H to C long range HMBC signal from hfOrn C-2 and hfOrn-NH to hOrn C-1 clearly confirmed the connectivity between hOrn and hfOrn residues. The location of the two formyl groups was also identified from HMBC signals between the formyl proton and C-5 in case of hfOrn and between C-5 proton and C-6 carbonyl carbon in the case of ahfOrn.

^{1}H-^{1}H through space correlations from NOESY and difference NOE spectra measured on the 700MHz NMR spectrometer were also obtained. Key correlations are summarised in the Figure 4-18. These correlations confirmed the location of the acetyl and formyl groups and the connectivity of the amino acid residues within foroxymithine. The combined NOESY/difference NOE data also provided evidence for the diketopiperazine and the connection between C-5 of hOrn and C-1 of Ser residue.

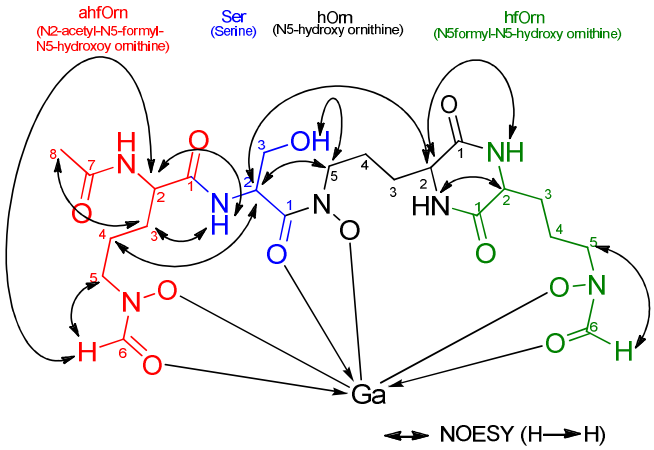
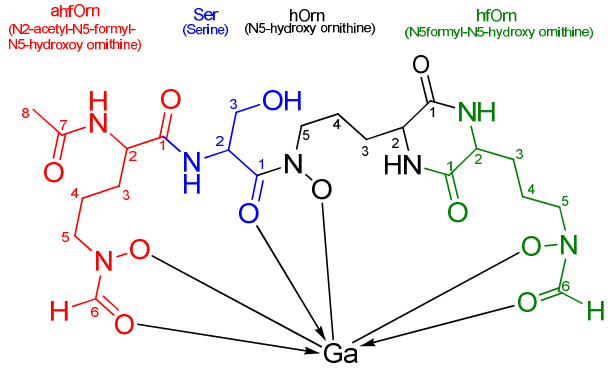


Figure 4-18 Key long range correlations observed in NOESY/NOE difference spectra of Ga-foroxymithine

Analysis of the combined data from the 1D and 2D NMR experiments resulted in ^{1}H and ^{13}C chemical shifts assignments for each H and C atom of Ga-foroxymithine (Table 4.3).

Table 4.3 ^{1}H and ^{13}C NMR assignments for Ga-foroxymithine in d6-DMSO (duplicate signals are given in brackets)

Assignment	δH	δC	COSY	HMBC
ahfOrn (N2-acetyl-N5-formyl-N5-hydroxy ornithine)				
ahfOrn-C1	-	171.13	-	
ahfOrn-C2	4.50 (4.47)	50.72	1.53, 1.35	18.20, 27.85, 171.13
ahfOrn-C3	1.35/1.53	27.85	3.35, 1.6	18.20, 27.85
ahfOrn-C4	1.60/1.38	18.28	1.53, 3.35	
ahfOrn-C5	3.35/3.47	50.33	1.60	27.85, 18.28, 152.3
ahfOrn-C6	-	168.7		-
ahfOrn-C7	-	169.0		-
ahfOrn-C8	1.84	22.4	-	169.0
ahfOrn-C(O)H	8.05 (8.07)	152.3	-	50.33
ahfOrn-NH	7.75 (7.79)	-		169.0, 50.72
Ser (Serine)				
Ser-C1	-	161.36		
Ser-C2	4.71 (4.66)	49.79	3.48, 5.06	161.36, 171.13
Ser-C3	3.48	60.3	4.71,	
Ser-NH	8.40	-		171.13
Ser-OH	5.06 (5.16)	-	3.48, 4.71	-
hOrn (N5-hydroxy ornithine)				
hOrn-C1	-	168.09	-	
hOrn-C2	3.58	54.64	1.51, 8.17	168.09
hOrn-C3	1.51/1.59	30.7	3.58, 1.79	
hOrn-C4	1.79/1.50	22.8	1.51, 3.45	
hOrn-C5	3.45	49.10	1.79	30.07
hOrn-NH	8.21 (8.17)	-	3.69	54.64, 168.09
hfOrn (N5-formyl-N5-hydroxy ornithine)				
hfOrn-C1	-	168.03		
hfOrn-C2	3.50	54.5	1.59	168.03, 31.9
hfOrn-C3	1.59/1.51	31.9	3.50, 1.48	
hfOrn-C4	1.48/1.81	22.4	1.59, 3.48	
hfOrn-C5	3.48/3.66	50.2	1.48	152.0
hfOrn-C(O)H	8.14	152.0		152.0
hfOrn-NH	8.24 (8.25)	-	3.50	54.5



To confirm the NMR structural assignment, ESI-MS/MS analysis was performed (Figure 4-19). Fragment ions observed were consistent with the proposed planar structure for foroxymithine.

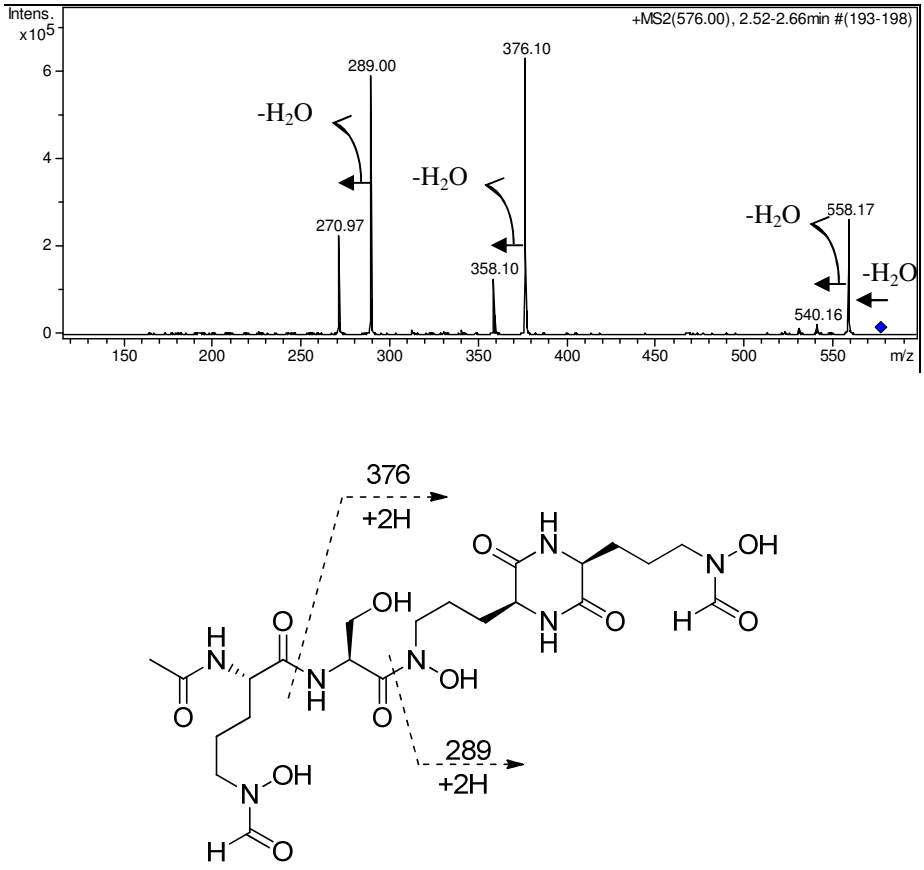


Figure 4-19 ESI-MS/MS analysis of foroxymithine

4.3 Stereochemical investigations of foroxymithine

We had originally planned to elucidate the relative stereochemistry of foroxymithine via a combination of internuclear distances derived from NOESY/NOE difference data and dihedral angles derived from ^{1}H NMR data. However the signal duplication observed in the ^{1}H NMR spectrum of Ga-foroxymithine rendered this approach impractical. Thus, the absolute stereochemistry of each of the amino acid constituents of Ga-foroxymithine was determined using Marfey's method. This method is based on hydrolysis of the peptide of interest using 6M HCl, followed by reaction of the products with 1-fluoro-2,4-dinitrophenyl-5-alanineamide (FDAA - also known as

Marfey's reagent). The resulting amino acid derivatives can be compared with authentic standards using LC-MS to determine their absolute configuration [238].

Hydrolysis of Ga-foroxymithine should yield N5-hydroxyornithine and serine residues. While D and L serine are both commercially available, D and L-N5-hydroxyornithine are not. Therefore, in order to apply Marfey's method to elucidate the stereochemistry of foroxymithine, authentic standards of D- and L-N5-hydroxyornithine needed to be synthesized.

4.3.1 Chemical synthesis of D- and L-N5-hydroxyornithine

Authentic standards of D- and L-N5-hydroxyornithine were chemically synthesised from commercially available D and L-ornithine via the route outlined in Figure 4-28.

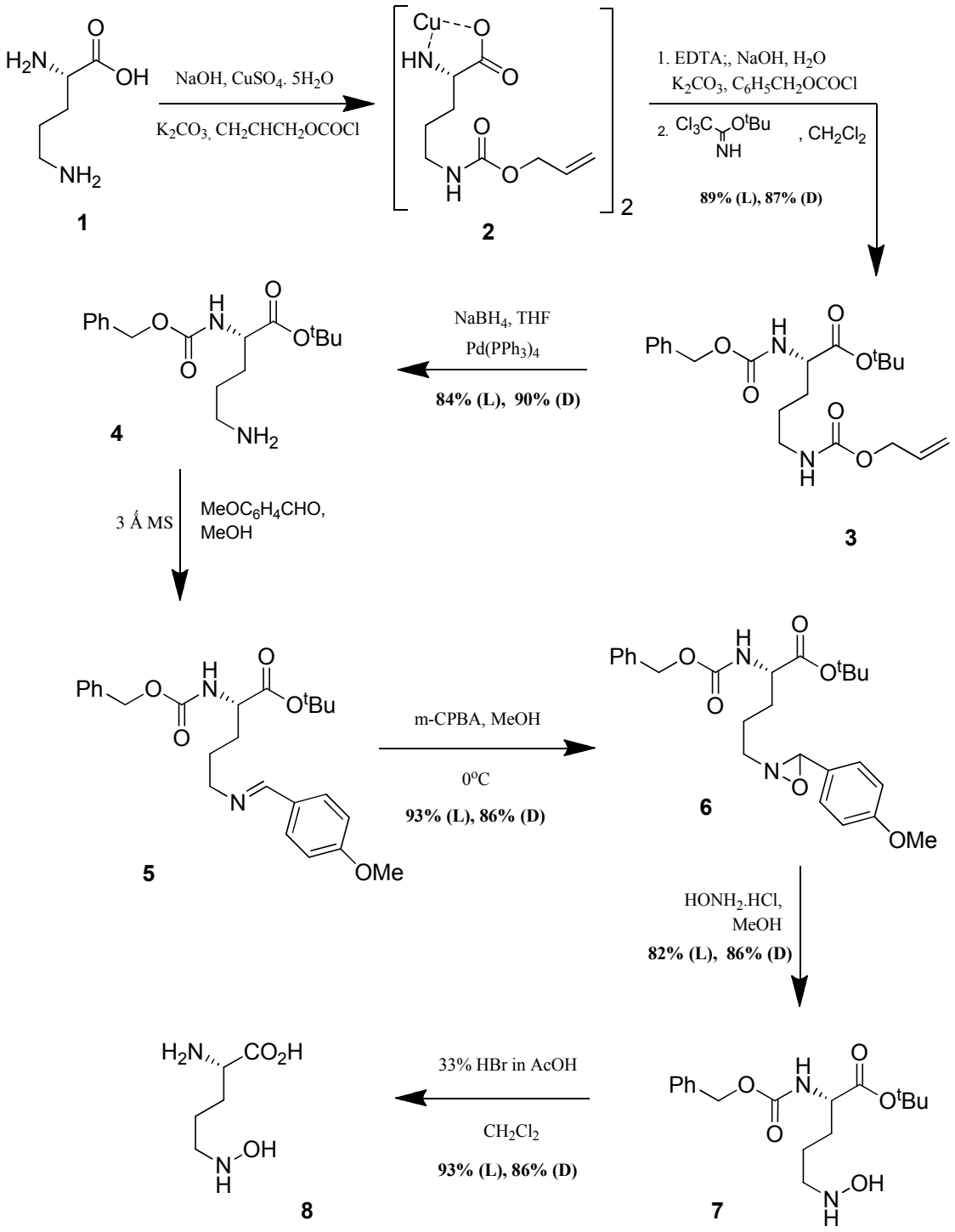


Figure 4-20 Reaction scheme for chemical synthesis of N5-hydroxyornithine from L-ornithine. An identical route was used for the synthesis of D-N5-hydroxyornithine from D-ornithine

The α-amino and carboxyl groups in ornithine (**1**) were protected under basic conditions via formation of a Cu^{2+} coordination complex. The α-amino group was then protected via addition of alloc chloride and potassium carbonate to the reaction

mixture. Cu^{2+} was removed from the crude product **2** by reacting it with EDTA and potassium carbonate, followed by benzyloxycarbonyl chloride was added to protect the α-amino group of the resulting intermediate. Addition of trichloroacetimidate to the mixture resulted in the crude t-butyl ester **3**, which was purified by flash column chromatography. The alloc protecting group was removed from **3** by reduction using sodium borohydride in the presence of catalytic Pd to yield crude **4** which was purified by flash column chromatography. Reaction of **4** with *p*-anisaldehyde produced the imine **5**, which was converted to oxaziridine **6** by reacting with *m*-chlorobenzoic acid in MeOH at 0^{o}C. The purified oxaziridine was converted to hydroxylamine **7** by reacting with the hydrochloride salt of hydroxylamine in dry methanol. Deprotection of purified **7** was effected by reaction with 33% HBr in AcOH to give the desired product N5-hydroxyornithine (**8**).

4.3.2 Application of Marfey's method to stereochemical elucidation of the amino acid residues within foroxymithine

The products resulting from reaction of a racemic mixture of D- and L-serine with Marfey's reagent are shown in Figure 3-28. This illustrates the process by which diastereomeric products, which are separable by chromatography, can be obtained from a racemic mixture of an amino acid.

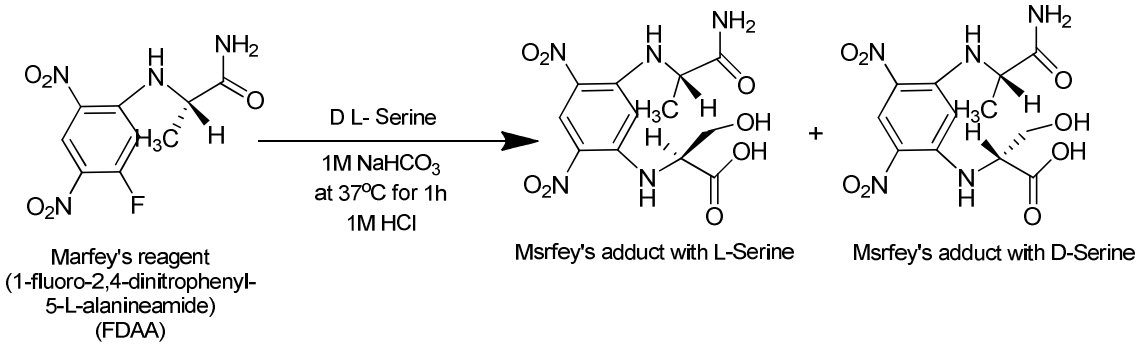


Figure 4-21 Reaction of Marfey's reagent with D L-serine to produce diasteroisomeric products that can be separated by chromatography

Stereochemical assignments were made by comparison of the retention times of Marfey's derivatives of authentic standards of D- and L-serine and D- and L-N5-

hydroxyornithine with Marfey's derivatives of the acid hydrolysis products of purified Ga-foroxymithine. The molecular formulae of the Marfey's derivatives of the authentic standards were confirmed by high resolution mass spectrometry (Table 4.4).

Table 4.4 High resolution MS analysis of Marfey's derivatives of authentic D- and L- serine and D- and L-N5hydroxyornithine

Compound	Measured *m/z*	Calculated *m/z*	Molecular formulae
Marfey's derivatised L-serine	358.9802 $(M + H)^+$ 380.0814 $(M + Na)^+$	358.0993 380.0813	$C_{12}H_{16}N_5O_8$ $C_{12}H_{15}N_5NaO_8$
Marfey's derivatised D-serine	358.0992 $(M + H)^+$ 380.0811 $(M + Na)^+$	358.0993 380.0813	$C_{12}H_{16}N_5O_8$ $C_{12}H_{15}N_5NaO_8$
Marfey's derivatised L-N5-hydroxyornithine	401.1411 $(M + H)^+$ 423.1234 $(M + Na)^+$	401.1415 423.1235	$C_{14}H_{21}N_6O_8$ $C_{14}H_{20}N_6NaO_8$
Marfey's derivatised D-N5-hydroxyornithine	401.1405 $(M + H)^+$	401.1415	$C_{14}H_{21}N_6O_8$

Figure 4-22 shows Extracted Ion Chromatograms (EIC) at *m/z* = 358 from LC-MS analysis of the Marfey's derivatised Ga-foroxymithine hydrolysate and the D- and L-serine Marfey's derivative standards. These chromatograms clearly showed peaks with distinctive retention times for the Marfey's derivatives of D and L serine. A peak in the EIC at *m/z* = 358 for the Marfey's derivatised Ga-=foroxymithine hydrolysate had the same retention time as the Marfey's derivative of L-serine. Co-injection of the Marfey's derivatives of the Ga-foroxymithine hydrolysate and D- or L-serine confirmed that foroxymithine contains L-serine.

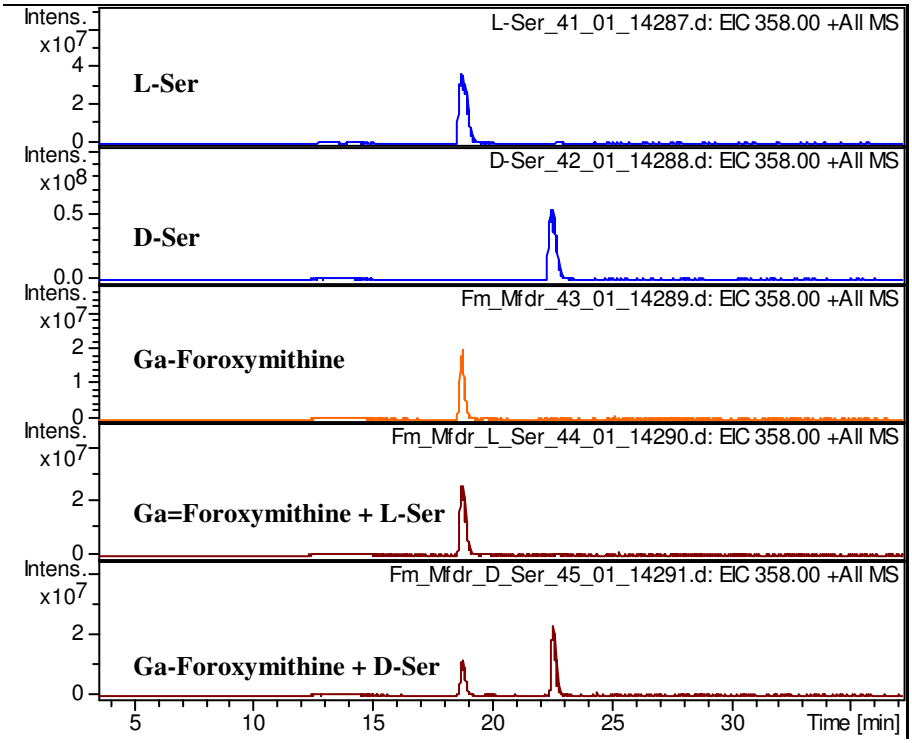


Figure 4-22 Comparison of Extracted Ion Chromatograms at *m/z* = 358 from LC-MS analyses of Marfey's derivatives of authentic D- and L-serine, and Ga-foroxymithine hydrolysate.

Similar analysis was performed using the chemically synthesised authentic standards of D- and L-N5-hydroxyornithine. The results are shown in figure 4-23 and confirm that foroxymithine contains L-N5-hydroxyornithine, but not D-N5-hydroxyornithine.

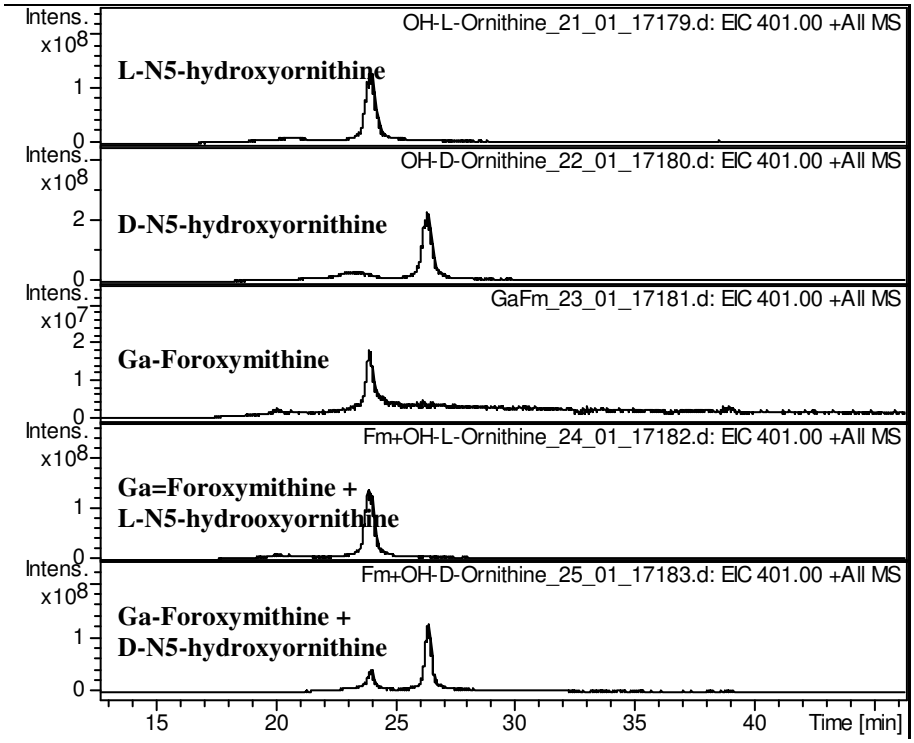


Figure 4-23 Comparison of Extracted Ion Chromatograms at *m/z* = 401 from LC-MS analyses of Marfey's derivatives of D- and L-N5-hydroxyornithine, and the hydrolysate of Ga-foroxymithine.

4.3.4 Application of Marfey's Method to determine the absolute stereochemistry of coelichelin

Although the relative stereochemistry of coelichelin has been deduced from NMR spectroscopic analysis of its gallium complex, the absolute stereochemistry of coelichelin remains unknown. It was therefore decided to apply Marfey's method to elucidate the absolute stereochemistry of coelichelin.

Coelichelin was purified from *Streptomyces coelicolor* M145 [83], using an identical protocol to that used for purification of foroxymithine from *S. narbonensis*. Purified Fe-coelichelin was hydrolysed and the hydrolysate was derivatised with Marfey's reagent. Subsequent LC-MS analyses, as described above for Ga-foroxymithine were performed on the Marfey's derivatised coelichelin hydrolysate and Marfey's derivatives of L-Thr, L-allo-Thr, D-Thr, D-allo-Thr (Figure 4-24). These analyses showed that coelichelin contains D-allo-Thr, as previously hypothesised [83].

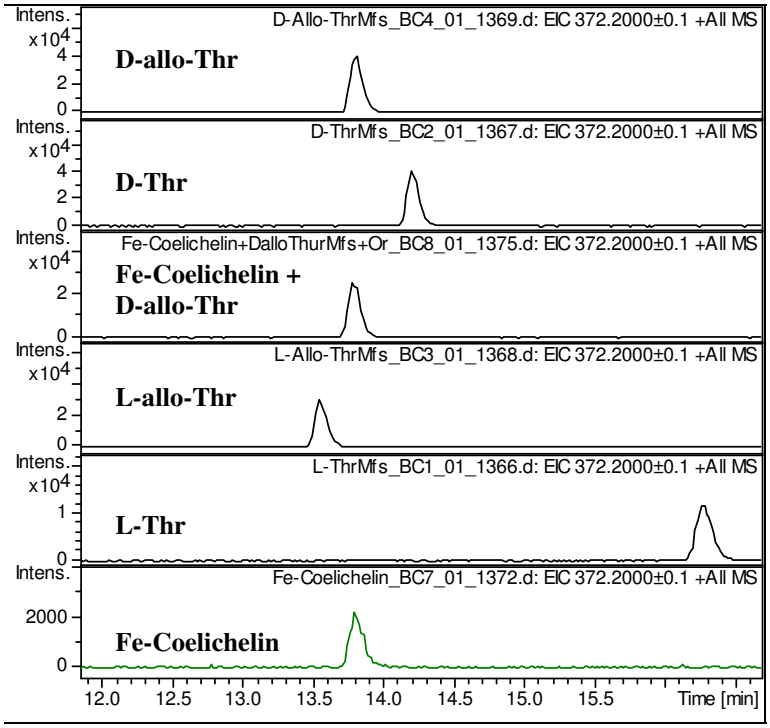


Figure 4-24 Comparison of Extracted Ion Chromatograms at *m/z* = 372 from LC-MS analyses of Marfey's derivatives of D-allo-Thr, D-Thr, L-allo-Thr, L-Thr and a coelichelin hydrolysate.

Similarly analyses were carried out using the Marfey's derivatives of chemically synthesised D- and L-N5-hydroxyornithine (Figure 4-25). However, the results were inconclusive.

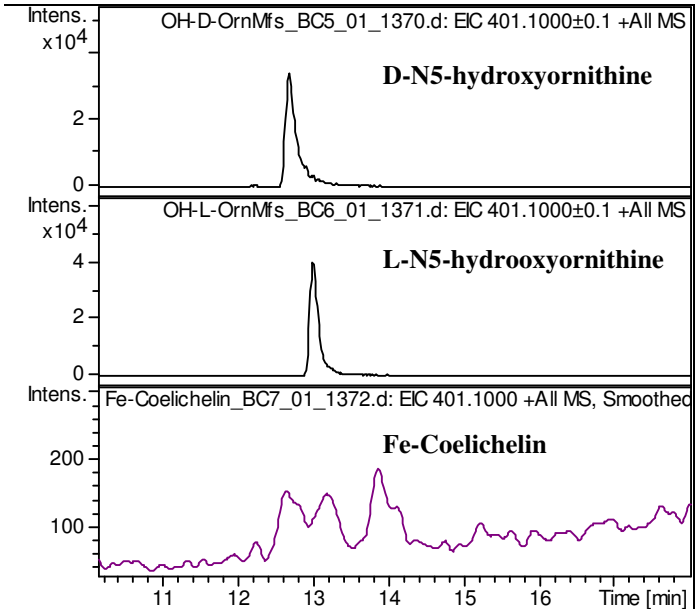


Figure 4-25 Comparison of Extracted Ion Chromatograms at *m/z* = 401 from LC-MS analyses of Marfey's derivatives of D-N5-hydrooxyornithine, L-N5-hydrooxyornithine and a coelichelin hydrolysate.

In conclusion, the application of Marfey's method has confirmed that foroxymithine contains L-serine and L-N5-hydroxyornithine, whereas coelichelin contains D-allo-threonine. Thus, despite the very similar planar structures of foroxymithine and coelichelin, the stereochemistry of these two tetrapeptides is dramatically different.

4.4 Attempted identification of the foroxymithine biosynthetic gene cluster in *S. narbonensis*

Foroxymithine and coelichelin have very similar structures. Both contain the non-proteinogenic amino acid N-formyl-N-hydroxyornithine. It has been shown that the *cch*A gene, encoding a putative formyl transferase, is involved in the biosynthesis of this amino acid (C. Corre, D. Oves-Costales and G.L. Challis unpublished data). The *cch*A gene lies within the coelichelin biosynthetic gene cluster as shown in figure 3-19.

We hypothesised that an orthologue of *cch*A is present within the foroxymithine biosynthetic gene cluster. Genomic DNA was isolated from *S. narbonensis* (the foroxymithine producer) using an established method [30]. Primers designed to anneal to conserved regions of *cch*A orthologues (kindly provided by Tingting Huang from the School of Life Science & Biotechnology, Shanghai Jiaotong University, Shanghai China) were used to amplify a fragment (~ 600 bp) of a putative formyl transferase encoding gene by PCR from *S. narbonensis* genomic DNA (Figure .4-26)

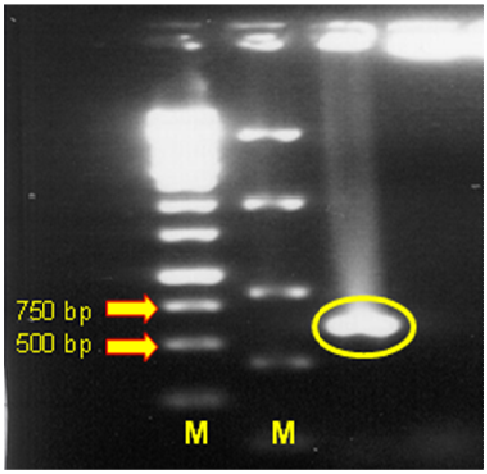


Figure 4-26 Agarose gel electrophoresis analysis of the amplification via PCR of a *cch*A orthologue from *S. narbonensis* genomic DNA. The lanes marked 'M' containing molecular size standards. The product of the reaction is circled in yellow.

The PCR product was purified and, after cloning into the pGEM T-Easy plasmid, it was sequenced. Alignment of the sequence obtained with the sequence of the *S. coelicolor* A3(2) corresponding fragment of the *cch*A gene showed that the sequences are >87% similar (Figure 4-27). This indicated that a *cch*A orthologue is present in *S. narbonensis.* This gene could lie within the foroxymithine biosynthetic gene cluster.

```
cchA      1 gcgggtcgtcatgttcggctaccagacctggggccaccgcaccctgcgagccctgctggactccgaacac
forA      1 gcgggtcgtcatgttcggttatcagacctggggccatcggaccctccgagcgctcctggactccgagcac

cchA     71 gacgtggtcctcgtggtcacgcatcccaggagcgagcacgcgtacgagaagatctggagcgactccgtcg
forA     71 gacgtggtgttggtcgtgacgcaccccaagagcgagcacgcctacgagaagatctggagcgactcggtcg

cchA    141 ccgacctcgccgaggagcacggcgtcccggtgctgatccgcaaccgcccggacgacgacgagctgttcga
forA    141 ccgatctggcggcggagcacggcgtccccgtcgtcatccgcaaccggcccgacgacgaggagctgttcga

cchA    211 gcgcctcaaggacgccgatccggacatcatcgtcgccaacaactggcggacctggatccccccgcgcatc
forA    211 gcgcctcaaggaggcggatccggacatcatcgtggccaacaactggcgtacgtggatcccgccgcgcatc

cchA    281 ttcggcctcccgcgccacggcacgctcaacgtgcacgactcgctgctgccgaagtacgccgggttctccc
forA    281 tacaccctgccgcgccacggcacgctcaacgtccacgactcgctgctgccgaagtacgccggattctcgc

cchA    351 cgctgatctgggcgctgatcaacggcgagaccgaggtgggcgtcaccgcgcacatgatgaacgacgagct
forA    351 ccctcatctgggcgttgatcaacggcgagcccgaagtgggcgtcaccgcacacatgatggacgaggtgct

cchA    421 ggacgccggtgacatcgtccgccaggaggccgtgccggtgggcccggccgacaccgccacggacctgttc
forA    421 cgacgccggcgacatcgtccggcaggagtcggtggccgtggagccgaccgacacggcgacggacctcttc

cchA    491 cacaagaccgtggacctcatcgccccggtcaccgtcggcgccctcgggctcatcgcctccgggcagacgg
forA    491 cacaagacggtcgacctcatcgccccggtcaccatcggcgccctcgacctgatcgcctccgggcagaccg

cchA    561 agttcaccaagcaggaccggtcccgggcgagcttcttccacaagcggtc
forA    561 agttcacgccacaggaccggtcccaggccaccttcttccacaagcggtc
```

Figure 4-27 Sequence alignment of the PCR product amplified from *S. narbonensis* genomic DNA with the corresponding fragment of the *S. coelicolor cch*A gene

Next, a fosmid library of *S.narbonensis* genomic DNA was prepared. The commercially available CopyControl™ Fosmid Library Production Kit (Epicenter® Biotechnologies) was used to prepare the fosmid library. An overview of the steps involved in library preparation is shown in Figure 4-28.

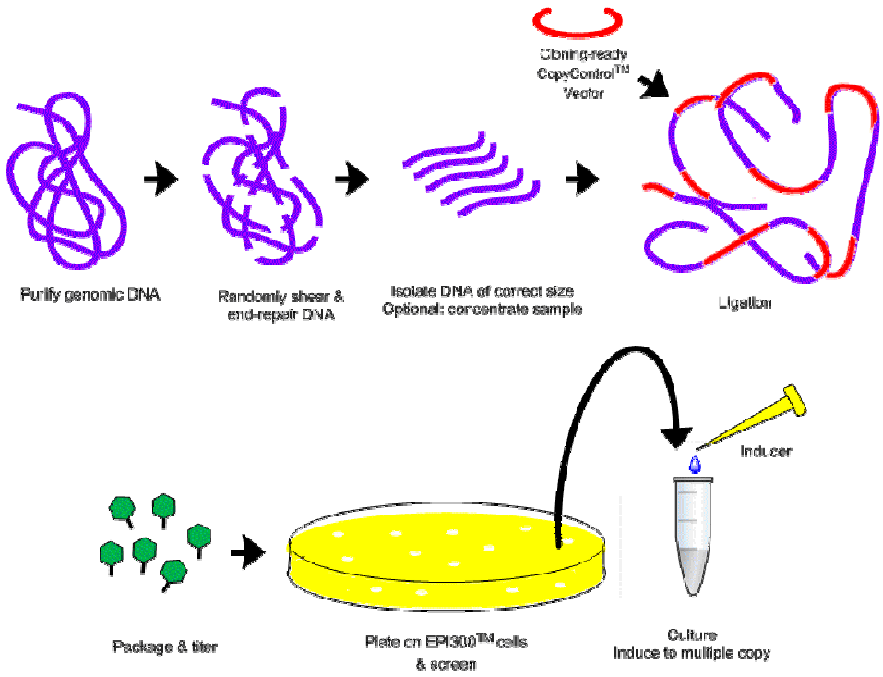


Figure 4-28 Overview of the process for preparing a fosmid library using the CopyControl™ Fosmid Library Production Kit

The process of fosmid library production involved shearing of the genomic DNA to yield fragments of approximately 40 kb in size. After end repairing the DNA, it was separated by overnight pulsed field gel electrophoresis. DNA of the desired size (~ 40 kb) was extracted from the gel and ligated with pCC1FOS (Figure 4-29). The ligation products were packaged into lambda phage using the MaxPlank lambda packaging extract supplied with the kit. *E. coli* EPI300 cells were infected with the phage and the cells were plated on LB agar plates containing chloramphenicol (12.5 µg/mL). After overnight incubation at 37°C, several chloramphenicol-resistant colonies had grown. A few of these were picked and inoculated into liquid LB medium. After overnight growth at 37°C, fosmid DNA was extracted from the cultures using the commercially available Epicentre Fosmid DNA extraction kit.

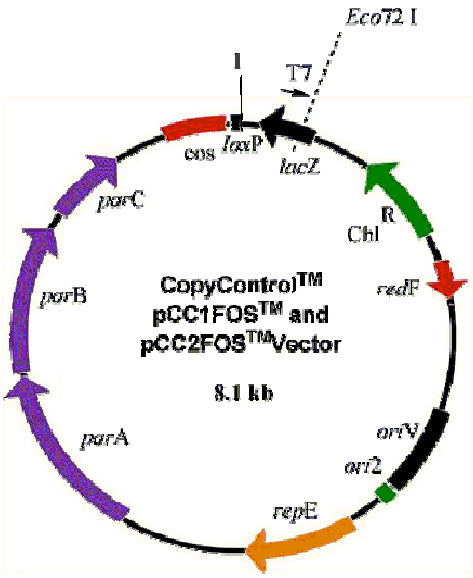


Figure 4-29 Map of the pCC1FOS fosmid vector

Before actually screening the library for clones containing the *cch*A orthologue amplified from *S. narbonensis* DNA by PCR, several control experiments were carried out on fosmid DNA isolated from randomly-picked clones, as follows.

a) PCR was performed on the following samples: genomic DNA, sheared DNA, and size-selected and end repaired DNA with the primers used to amplify the putative formyl transferase gene fragment from *S. narbonensis*. PCR products were analysed by agarose gel electrophoresis (Figure 4-30) and bands of the expected size were cut from the gel and purified using a commercially available Gel Extraction kit. The resulting DNA molecules were sequenced and the sequences obtained were analysed using BLAST. The results showed that the sequence was highly similar (90 % identical) to the *cch*A gene of *Streptomyces coelicolor*.

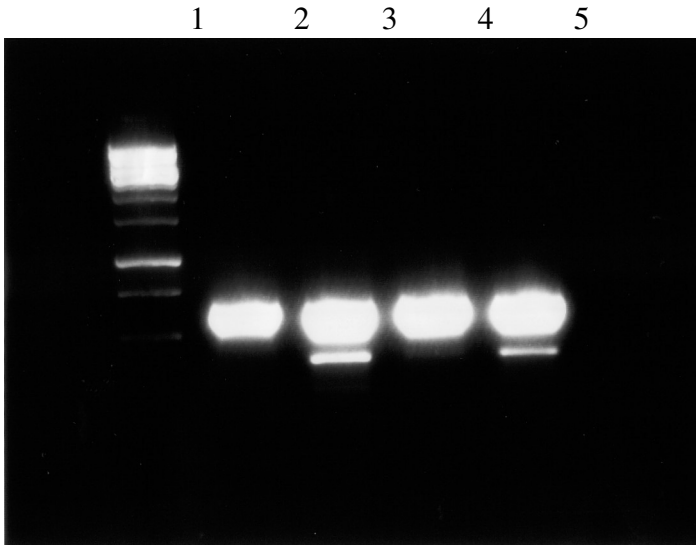


Figure 4-30 Agarose gel electrophoresis analysis of control PCR reactions (lane 1 = 1Kb DNA ladder, 2 and 5 = genomic DNA, 3 = Sheared DNA, 4 = Size selected & end repairedDNA)

b) PCR was performed on fosmid DNA isolated from randomly-selected clones using a set of primers called "fosmid primers" designed to amplify a region of 459bp of the pCC1FOS vector, and a set of primers called "pCC1FOS sequencing primers" designed to span the cloned insert within each fosmid. No product was obtained using the PCC1FOS sequencing primers, indicating that the clones contain large inserts (Figure 4-31). A product of approximately 459bp was obtained using the fosmid primers, indicating that the isolated DNA contained the pCC1FOS vector backbone (Figure 4-31). Together, the results showed that each randomly-picked clone contains the pCC1FOS vector with a large insert.

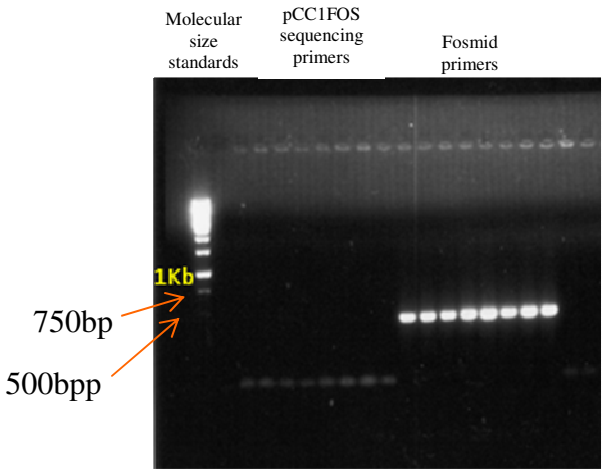


Figure 4-31 Agarose gel electrophoresis analysis of PCR products obtained from DNA purified from randomly picked clones

c) Fosmid DNA, purified from the randomly picked clones was digested with the *Bam*H1 restriction enzyme and analysed on a 0.7% agarose gel. Every clone should show a different digestion pattern if the library is derived from random segments of *S. narbonensis* genomic DNA. *Bam*H1 cuts the vector twice giving rise to bands of 8.1kb and 12bp. Thus an 8.1kb band would be expected for all clones. In the event, all of the fosmids purified from the randomly-picked clones yielded a digestion product of 8.1kb, along with other digestion products of random size (Figure 4-32).

Thus, the above sets of control experiments established that the fosmid library contains random ~ 40kb pieces of *S. narbonensis* genomic DNA and at least one clone that has the gene of interest within its insert.

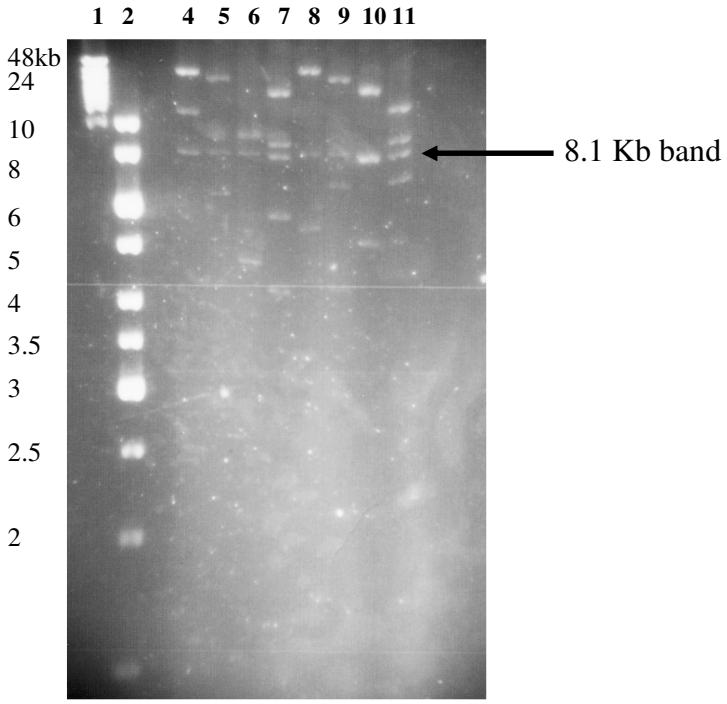


Figure 4-32 Agarose DNA gel electrophoresis analysis of *Bam*H1 digested fosmids isolated from randomly-picked clones from the S. *narbonensis* genomic library (lane 1 & 2 = DNA molecular size standards, lanes 4-11 = digested fosmids from randomly-picked clones)

4.4.1 Fosmid library screening

E. coli cells containing the fosmid library were plated on LB agar containing chloramphenicol and grown overnight at 37°C. The resulting colonies were picked separately into 96 well plates containing LB liquid medium + chloramphenicol. The

96 well plates were incubated overnight at 37°C and a few microlitres of culture from each of the wells of a single 96 well plate were combined and the fosmid DNA was extracted. The extracted DNA from each 96 well plate was screened by PCR using specific primers for the putative formyl transferase gene fragment identified in *S. narbonensis*. Combined fosmids from 16 separate 96 well plates were screened and one of these plates yielded a product of the expected size for the S. *narbonensis* putative formyl transferease gene fragment (Figure 4-33).

Each well of the 96 well plate containing the hit was separately screened by PCR, but the well containing the fosmid of interest could not be identified. Due to time constraints it was not possible to pursue the identification, cloning and sequencing of the putative foroxymithine biosynthetic gene cluster in *S. narbonensis* further.

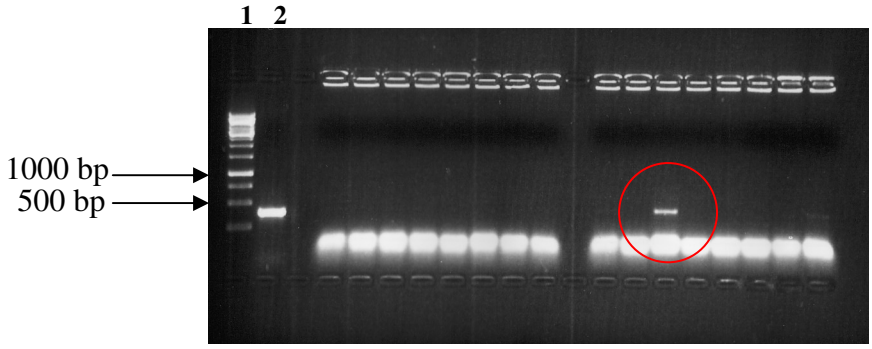


Figure 4-33 Agarose gel showing a positive hit during *S. narbonensis* genomic DNA fosmid library screening from PCR using primers targeting *cch*A homologue gene (lane 1 = 1Kb DNA molecular size standards, 2 = positive control using *S. narbonansis* genomic DNA as template, the rest of the lanes contain fosmid DNA samples from different 96 well plates)

4.5 Conclusions and future work

The overall aim of this part of the project was to carry out structural and biosynthetic investigations of foroxymithine. Foroxymithine was purified from a novel source (*S. narbonensis*) after growth in an iron-deficient medium. Through a series of HPLC purifications, the Ga-foroxymithine complex was isolated, facilitating NMR spectroscopic analysis. Several 1- and 2-D-NMR experiments including ^{1}H, ^{13}C, COSY, HMBC, HSQC and NOESY were performed. These allowed assignment of all the ^{1}H and ^{13}C NMR signals for Ga-foroxymithine, confirming its structure. Next, the stereochemistry of foroxymithine was investigated. Due to signal doubling in the ^{1}H NMR spectrum of Ga-foroxymithine, resulting from two distinct low energy conformations (or possibly configurations) we were not able to use molecular modelling to elucidate the relative stereochemistry of the complex (as had been done previously for coelichelin). Thus, the absolute stereochemistry of each of the amino acid residues in foroxymithine was determined using Marfey's method. In one experiment, Marfey's derivatives of D and L-serine residues were used as authentic standards to compare using LC-MS with Marfey's derivatised Ga-foroxymithine hydrolysate. The results showed that foroxymithine contains L-serine. An authentic standard of N5-hydroxyornithine was not commercially available. Thus, D- and L-N5-hydroxyornithine were synthesised and subjected to Marfey's derivatisation. These derivatives were compared with the Marfey's derivatised Ga-foroxymithine hydrolysate using LC-MS. The results demonstrated that foroxymithine contains only L-N5-hydroxyornithine. As the absolute stereochemistry of coelichelin has not been reported, it was purified from *S. coelicolor* M145 as its ferric complex. This was hydrolysed and derivatised using Marfey's reagent, and compared with Marfey's derivatives of L-Thr, D-Thr, L-allo-Thr and D-allo-Thr using LC-MS. This showed that the absolute stereochemistry of the threonine residue in coelichelin is D-allo. Thus, although there is high structural similarity between foroxymithine and coelichelin, they possess radically different stereochemistries.

Foroxymithine contains two N-formyl groups. Thus to identify the putative foroxymithine biosynthetic gene cluster it was decided to design primers to target putative formyl tranferase genes in *S. narbonensis*. PCR using these primers and *S.*

narbonensis genomic DNA as a template produced a DNA fragment of 600bp, which upon sequence analysis was found to be 90% identical to the corresponding fragment of the *cch*A gene, which encodes a formyl transferase involved in coelichelin biosynthesis. A genomic fosmid library of *S. narbonensis* was constructed using a commercially available kit. Fosmid clones were screened by PCR with the primers used to identify the putative formyl tranferase gene fragment in *S. narbonensis*. However, attempts to identify a fosmid clone containing the formyl transferease gene fragment proved unsuccessful.

Future work may involve constructing a fresh fosmid library to locate the foroxymithine biosynthetic gene cluster. Once the fosmid clone is identified through screening the library by PCR, targeting putative formyl transferase gene, the fosmid can be sequenced. The sequence data can be analysed to investigate the modular arrangement within the biosynthetic gene cluster of foroxymithine in *S. narbonensis*. Further investigation may include gene inactivation and knockout experiments to confirm and identify the role of various enzymes being encoded within the gene cluster.

CHAPTER 5 - EXPERIMENTAL

5.1 Instrumentation

HPLC analyses were performed using an Agilent 1100 instrument equipped with a binary pump and a diode array detector. For analysis of HPLC fractions by direct infusion a Bruker MicroTOF ESI-MS mass spectrometer was used in positive mode. PCR was performed using an Eppendorf Mastercycler Personal. DNA concentrations were measured using a nano-drop ND-1000 spectrophotometer. A Bio-Rad Power PAC300 was used for agarose gel electroporesis. A Bio-Rad CHEF MAPPER was used for Pulse Field Gel electrophoresis. A basaire laminar flow hood was used for handling *Streptomyces* strains. NMR analyses were performed on Bruker 400 or 700 MHz spectrometers. The latter was equipped with a TCI cryoprobe.

For small scale centrifugation a bench top Eppendorf centrifuge 5424 was used. For large volumes Eppendorf 5810R and Sorvall RC6 Plus centrifuges were used. An Innova 44 series incubator shaker was used for growing strains in liquid media. A Priorclave autoclave was used for sterilising growth media and glassware by autoclaving at 121^{o}C for 20 min.

For semi-preparative purification on HPLC reversed-phase HPLC Agilent Zorbax C-18 column, 21.2 x 10.0mm, 5µm particle size was used.

Analytical column on LC-MS used was Agilent Eclipse XDB-C18, 4.6 x 150mm, particle size 5µ.

For stereochemical studies on Ga-foroxymithine complex two chiral columns manufactured by Daicel Chemical Industries were used, Chiralpak® IA 0.46cm x 25cm and Chiralpak® IC 0.46cm x 25cm.

CD spectra were recorded on a Jasco J-815 CD spectrometer.

5.2 Reagents, chemicals and bacterial growth media

All chemicals and reagents used in the study were of AnalR or molecular biology grade and were obtained from Sigma Aldrich (USA), Difco (USA), Fisher Scientific (UK), Acros Organics (Belgium), Invitrogen (USA), VWR, Becton, Oxoid or Dickinson and Co or otherwise stated.

Restriction enzymes, high fidelity polymerase and Taq Polymerase were provided by MBI Fermentas (Lithuania), New England Biolabs (USA) or Roche (Germany).

Primers for PCR were ordered from Sigma Genosys (USA).

5.2.1 Buffers and general solutions

Stock solutions were prepared using the methods of Sambrook and Russell [251]. All solutions were made in water unless stated otherwise

Table 5.1 Buffers and general solutions

Name	Chemical composition/stock solutions
Tris-HCl (pH8)	0.5M (pH adjusted with HCl)
TE buffer	100mM Tris-HCl pH8, 10mM EDTA pH8
TBE 5x stock solution	For 1L 53g Tris base, 27.5g boric acid, 10mM EDTA pH8
Loading buffer for agarose gels	50% glycerol, 0.1% bromophenol blue, 0.1% xylene cyanol
SET buffer	75mM NaCl, 75mM EDTA pH 8, 20mM Tris HCl pH 7.5
EDTA pH 8	0.5M
Tris-HCl pH 7.5	0.5M
NaCl	5M
Chloramphenicol stock	25mg/mL (in ethanol)
$NaHCO_3$	1M
HCl	6M
Buffer A (for LC-MS analysis of Marfey's derivatives)	10mM ammonium formate, 1% methanol, 5% acetonitrile, pH 5.2
Buffer B (for LC-MS analysis of Marfey's derivatives)	10mM ammonium formate, 1% methanol, 60% acetonitrile, pH 5.2
$FeCl_3$	100mM
H_2SO_4	0.5M

5.2.2 Growth Media

5.2.2.1 SMM (Supplemented Minimal Medium)

Table 5.2 SMM medium recipe

Ingredients	Quantity (mL)
Distilled water	819
$MgSO_4$. $7H_2O$ (24g/L)(0.2M)	25
TES buffer (0.25M, pH 7.2)	100
NaH_2PO_4 (6g/L)+ K_2HPO_4 (6.8g/L)(50mM each)	10
Trace element solution	1
Casaminoacids (20% w/v)	10
Total	1000 mL
Then Glucose (20%) 25mL	50

5.2.2.2 Starch Casein Agar medium for *Streptomyces griseus*

Following ingredients were added to 1L distilled water and autoclaved.

Table 5.3 Starch Casein Agar medium recipe

Ingredients	Quantity
Soluble starch	10g
Vitamin-free Casein	0.3g
KNO_3	2.0g
NaCl	2.0g
K_2HPO_4	2.0g
$MgSO_4$ $7H_2O$	0.05g
$CaCO_3$	0.02g
$FeSO_4$. $7H_2O$	0.01g
Agar	20.0g

5.2.2.3 Iron Deficient Medium [83]

To prepare iron deficient medium 2 g of K_2SO_4, 3 g of K_2HPO_4, 1 g of NaCl and 5 g of NH_4Cl were added into 1 litre distilled water. The mixture was then stirred with 50 g of Chelex-100 (Na form) overnight. Chelex resin was filtered off using Whatmann 1 filter paper. Trace elements were then added which included, 100µL of thiamine stock solution (20 mg/ml), 100 µL of $ZnSO_4.7H_2O$ stock solution (20mg/ml), 10 µL of $CuSO_4$ stock (0.5 mg/ml), 10 µL of $MnSO_4$ H_2O stock (3.5 mg/mL) and 80 mg of $MgSO_4.7H_2O$. Mixture was then autoclaved and before starting 50 mL culture in flask

500 µL each of the stock solution of $CaCl_2$ (10 g/L), glucose (250 g/L) and yeast extract (0.5%) were added.

5.2.2.4 TSB (Tryptone soy Broth)

To prepare 1 litre of TSB culture 30 g of Oxoid Tryptone soy Broth powder (30 g) was used. After mixing it was then autoclaved.

5.2.2.5 LB Agar medium

To prepare 1 litre of LB agar 25 g of Luria-Bertani powder and 15 g of Difco Bacto Agar were mixed together and autoclaved.

5.2.2.6 Soy Flour Mannitol Medium (SFM) [30]

To prepare 1 litre of SFM 20 g of Soya flour, 20 g of D-mannitol and 15 g of Bacto agar were dissolved in tap water and made up the volume to 1 litre and autoclaved.

5.2.2.7 YEME medium

To prepare the medium 3 g of Difco yeast extract, 5 g of Difco Bacto peptone, 3 g of Oxoid malt extract, 10 g of glucose and 340 g of sucrose were dissolved in distilled water and the volume was made up to 1 litre. After autoclaving, 2 mL of 2.5 M $MgCl_2.6H_2O$ stock was added (5 mM final) and finally 25 mL of 20% glycine stock (0.5% final) was added.

5.2.3 Primers

Following is the list of primers used in the study.

Table 5.4 List of primers

Primer name	Primer sequence (5' – 3')	Purpose
165-CR-FOR	AAGTCGAACGATGAAGCC	16S-RNA sequencing primers
165-CR-REV	ACCGTCGAACATTTGCCA	
16S-Cl-For	CCGAATTCGTCGACAACAGAGTTTGATC CTGGCTCAG	
16S-Cl-Rev	CCCGGGATCCAAGCTTAAGGAGGTGATC CAGCC	
pCC1FOS-Fw	GGATGTGCTGCAAGGCGATTAAGTTGG	pCC1FOS sequencing primers
pCC1FOS-Re	CTCGTATGTTGTGTGGAATTGTGAGC	
forA-Fw	TTGCGGGTCGTCATGTTCG	Fosmid library screening primers to amplify a formyl transferase homologue gene in *S. narbonensis*
forA-Re	GACCGCTTGTGGAAGAAGG	
cchA-Fw	TGCGGGTCGTCATGTTC	
cchA-Re	GACCGCTTGTGGAAGAAGG	
pm1-Fw	GGCGTGAACTCGGTCTGCCC	
pm1-Re	CCTCCGAGCGCTCCTGGACT	
pm7-Fw	AAGAGGTCCGTCGCCGTGTC	
pm7-Re	AGAGCGAGCACGCCTACGAG	
fmft-Fw	TGCGGGTCGTCATGTTCG	
fmft-Re	GACCGCTTGTGGAAGAAGG	

5.3 Bacterial strains

The bacterial strains used in this research and their origin are listed in table 4-1, below.

Table 5.5 List of Bacterial Strains used in the study

Name	Purpose	source
S. griseus DSM 4023	Fish toxins purification	DSMZ (Germany)
Actino 13 (*S. griseus*)	Fish toxins purification	Dr. Jackie Parry, Lancaster University
Actino 2 (*S. griseus*)	Fish toxins purification	Dr. Jackie Parry, Lancaster University
Actino 6 (*S. griseus*)	Fish toxins purification	Dr. Jackie Parry, Lancaster University
Actino 12 (*S griseus*)	Fish toxins purification	Dr. Jackie Parry, Lancaster University
Strain 82 (*S. griseus*)	Fish toxins purification	Dr. Jackie Parry, Lancaster University
Strain 62 (*S. griseus*)	Fish toxins purification	Dr. Jackie Parry, Lancaster University

S. narbonensis DSM 40016	Foroxymithine purification	DSMZ (Germany)
S. narbonensis G3P3/110	Foroxymithine purification	Lab stock
E.coli EPI300™-TI®	Fosmid library preparation	Epicentre® Biotechnologies

5.3.1 Production and storage of bacterial strains

Working under sterile conditions *Streptomyces griseus* and *Streptomyces narbonensis* spore stocks were made by adding 10-20 µL of spores to 100 µL of water and applied on SFM plate followed by incubation at 30°C for 7-8 days. Sterile distilled water (3 mL) was added and using sterile spreader scratched off bacterial colonies lightly and thoroughly. Mycelial fragments were removed by filtration using a sterile syringe containing sterile glass wool. Filtrate was then centrifuged at 3000 rpm for 5 minutes and supernatant was discarded. 200 µL of sterile distilled water and 200 µL of sterile 50% glycerol was then added. After mixing contents were then transferred into a small sterile Eppendorf tube for storage at -20°C.

5.4 Sample purification methods

5.4.1 Fish toxins sample purification

Streptomyces griseus strains were inoculated on sterile cellophane membranes on starch casein agar medium in a Petri plate and incubated at 30°C for 7 days. Biomass was collected and weighed. After removing the membrane the agar was divided into two halves. Biomass and one half of the agar were soaked into 50 mL of distilled water and the second half of agar into 50mL of ethyl acetate. Agar was crushed using spatula into small pieces. The mixtures were left at 4°C for two days before filtering off the biomass and agar. The filtrate left from biomass filtration was identified as ‘aqueous filtrate’, ethyl acetate filtrate as ‘organic filtrate’, while the filtrate from the agar soaked into water was then subjected to organic extraction using ethyl acetate. After extraction the organic phase was identified as ‘organic extract’ and aqueous phase as ‘aqueous extract’. Organic filtrate and organic extract were dried under vacuum while aqueous fractions were frozen in liquid nitrogen and lyophilized on a freeze dryer.

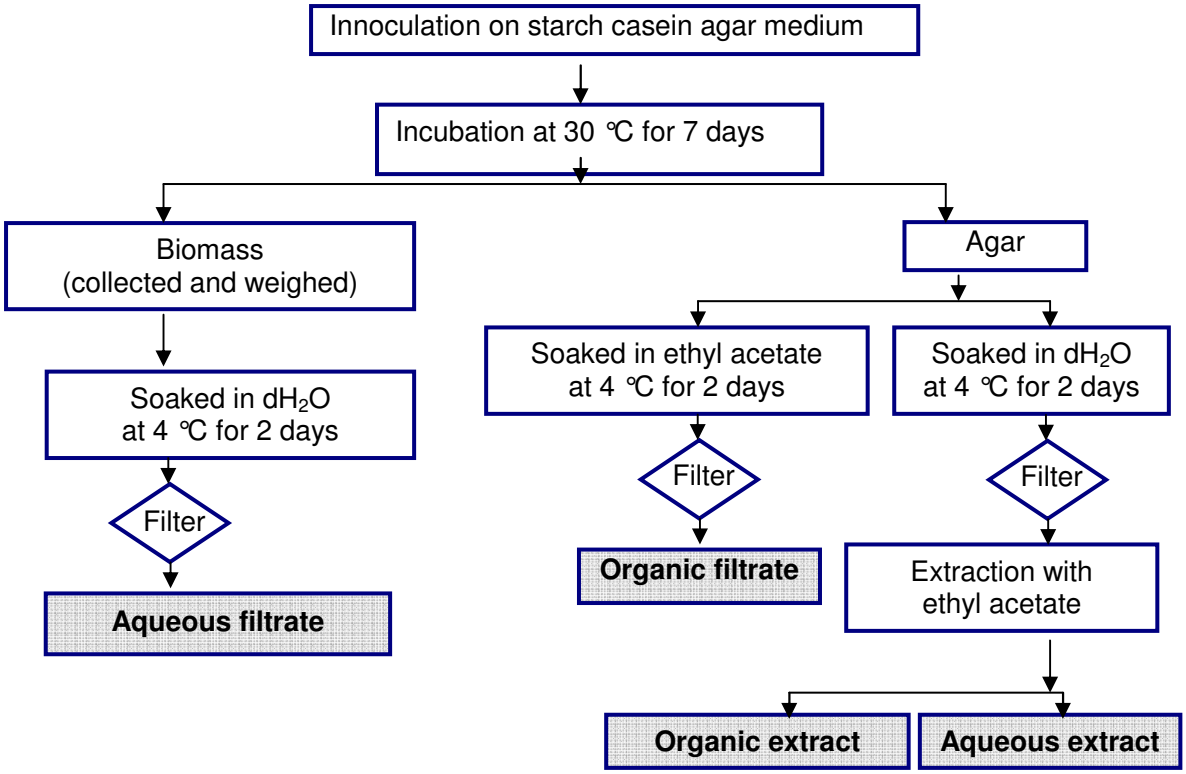


Figure 5-1 General scheme for fish-toxin sample preparation and purification

5.4.1.1 Sample preparation for fish toxin experiment

Based on the weight of the biomass, calculations were made to dissolve dried stock samples into specific quantity of solvent (water in case of aqueous sample and ethyl acetate in case of organic samples) to prepare three dilutions of the samples, I) equivalent to 6mg of biomass, II) 12mg of biomass and III) 18mg of biomass. Required quantities of stock samples were transferred into glass vials and dried (organic samples under vacuum and aqueous samples on freeze dryer). When dry they were shipped for fish-exposure experiments in London.

5.4.2 Foroxymithine purification

5.4.2.1 Growth of *Streptomyces narbonensis* in iron deficient medium

For foroxymithine purification *Streptomyces narbonensis* G3P3/110 and *Streptomyces nabonensis* type strain DSM 40016 were grown in 500mL iron deficient medium with an inoculation of 80-100 µL of spore stock. Flasks were then incubated on rotary shaker at 30°C for 4-5 days.

5.4.2.2 Preliminary purification on HP-20 resin

Iron deficient medium was centrifuged after incubation at 3000 rpm for 10 minutes. Spent media were filtered through Whatman No. 1 filter paper. TFA was added into spent media at final conc. of 0.05% (i.e. 0.05mL or 50µL in 100mL). HP-20 resin was put into a glass chromatography column. The column was eluted with 100% MeOH (3 volumes of resin). Solvent was then replaced with distilled water containing 0.05% TFA gradually by adding and the resin was washed with 3 volume of distilled water containing 0.05%TFA. Spent media containing 0.05%TFA was loaded. It was then washed with distilled water containing 0.05%TFA (3 volumes of resin). Finally 20% MeOH (no TFA) was used to elute foroxymithine. After freezing in liquid nitrogen the fraction containing forxymithine was subjected to freeze drying.

5.4.2.3 Purification of Ferric-foroxymithine complex by HPLC

Freeze-dried sample was then dissolved in $FeCl_3$ to a final concentration of 10 mM (of $FeCl_3$). The solution turned dark red. The mixture was then centrifuged at 13,000 rpm for 5 minutes and the supernatant was separated to be injected on HPLC (Agilent 100) for purification.

UV detector on HPLC was set to show absorbance at 435 nm and flow rate was set at 5 mL/min using C18 RP column. The method set was as below

Table 5.6 HPLC method for ferri-foroxymithine purification

Time	A (%)	B (%)
0	100	0
20	90	10
25	0	100
35	0	100
40	100	0
50	100	0
Flow rate 5mL/min, UV = 435nm A= water + 0.1% TFA; B= acetonitrile + 0.1% TFA		

Peak fractions were analysed on Bruker microTOF ESI-MS to see if they contained ferri-foroxymithine. The retention time of ferri-foroxymithine was around 10 minutes. Collected fraction were put together in a round bottom flask and after putting on rotary vacuum for 15-20 minutes to remove any acetonitrile left, put on freeze dryer after freezing in liquid nitrogen.

5.4.2.4 Removal of Iron from ferri-foroxymithine

All glassware used were rinsed with 6N HNO_3 and MilliQ water. Using an established protocol[249] freeze-dried ferri-foroxymithine dry powder was solubilised in 25 mL deionized water. To this solution 145 mg (1 mmole) of 8-hydroxyquinoline dissolved in 25 mL MeOH was added. Solution turned from yellowish-red to dark green as Fe(II) complexed to 8-hydroxyquinoline. It was left in shaker for 1 hour at room temperature. Fe-quinoline complex was then extracted by using 20 mL DCM (dichloromethane). Desferri-foroxymithine was in aqueous phase. The extraction was repeated five times. Excess organic solvents were evaporated under vacuum. The flask was snap-frozen in liquid N_2 and lyophilized overnight.

5.4.2.5 Preparation and purification of gallium-foroxymithine complex

The freeze-dried sample of desferri-foroxymithine was dissolved in 5 mL of deionised water and mixed with the gallium sulphate solution (0.151 g of gallium sulphate in 5 mL of 0.05 M sulphuric acid). The solution was left for 30 minutes on stirrer. After 30 minutes pH neutralized and solution was frozen in liquid nitrogen and put on freeze-dryer. Freeze-dried gallium-foroxymithine complex was diluted into 5 mL of deionised water. Supernatant was separated after centrifugation for 8 minutes at 3000 rpm.

The gallium-foroxymithine complex was separated by RP-HPLC using a C18 RP semi-preparative column. The column was eluted with 100% 10 mM ammonium hydrogen carbonate buffer (pH 7), under isocratic conditions. 500 µL of diluted sample of gallium-foroxymithine was injected and eluted for 30 minutes with flow

rate of 5mL/min and detector wavelength set at 210 nm. The peak at approx 4-5 minutes was collected and freeze-dried.

5.4.3 Genomic DNA purification

Streptomyces narbonensis strain spore stock (10 µL) was inoculated in 25 mL TSB medium in 250 mL baffled conical flask. The flask was incubated on shaker (180 rpm) for 4 days at 30°C. After incubation 1 volume of water (25 mL) was added and then centrifuged for 10 min. at 2000 rpm. The tube was then washed in 50 mL of EDTA (10mM pH 8.0) and centrifuged again for 10 min. at 2000 rpm. Mycelia in the pellet were then resuspended in 5mL SET buffer and 100 µL lysozyme aqueous solution (50mg/mL) was added to it and then incubated for 1h at 37°C. 140 µL of proteinase K aqueous solution (20 mg/mL) was then added followed by 600 µL 10% SDS and mixed by inversion. It was incubated for 2 h at 55ºC. After incubation 2 mL of 5 M NaCl was added and mixed thoroughly by inversion, let to cool to 37ºC and then 5mL chloroform was added and mixed by inversion for 30 min at RT. It was centrifuged for 15 min at 4000 rpm. Supernatant was transferred to fresh 15 mL tube (about 6mL) then added 0.6 vol. isopropanol (about 3.6 mL), mixed by inversion, after about 3 min. DNA was fished out with plastic loop, rinsed in 5 mL 70% ethanol and was transferred in a 1.5 mL eppendorf to dry for 30 min. It was then resolubilised in 1 mL TE (pH 8.0) O/N at 4°C, then prepared 300 µL aliquots and stored at -20°C.

5.5 Fosmid library preparation

Fosmid library was prepared using commercially available 'CopyControl™ Fosmid Library Purification' from EPICENTRE Biotechnologies with the following protocol:

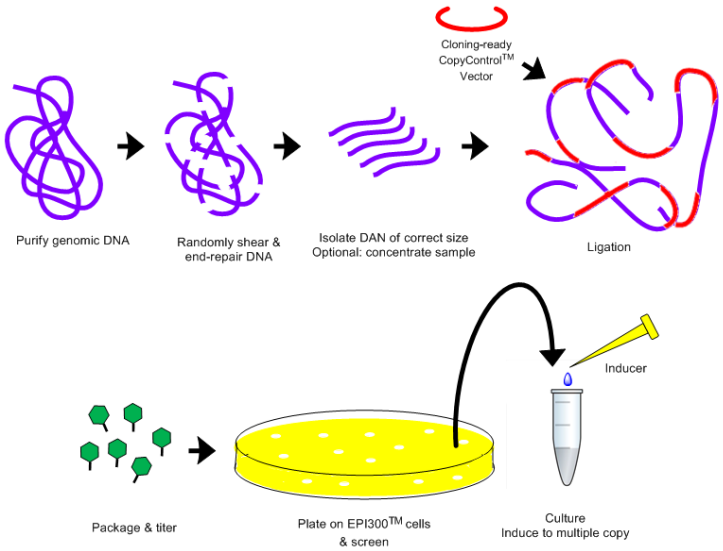


Figure 5-2 Overview of the CopyControl™ Cloning System[252]

Following steps are involved in the fosmid library preparation (detail of each step can be found in the manual (CopyControl™ HTP Fosmid Library Production Kit, Catalogue number CCFOS059).

1. Genomic DNA from liquid culture of *Streptomyces narbonensis* was purified using protocol (reference section 5.5.3)
2. DNA was sheared to approximately 40 Kb fragments.
3. Sheared DNA was end-repaired to blunt, 5'-phosphorylated ends.
4. End-repaired DNA was size resolved by LMP (Low Melting Point) agarose gel electrophoresis.
5. Blunt-end DNA purified from the LMP agarose gel.
6. Purified blunt-end DNA ligated to the Cloning-Ready CopyControl™ pCC2FOS™ vector.
7. Ligated DNA was packaged and plated on EPI300-T1® plating cells. Clones were grown overnight at 37ºC.
8. Fosmid clones were grown in LB media containing chloramphenicol and screened by PCR using specific primers.
9. Clones of interest were induced to high copy number using the CopyControl Induction Solution.
10. Fosmid DNA was purified using QIAprep miniprep kit from Qiagen for sequencing.

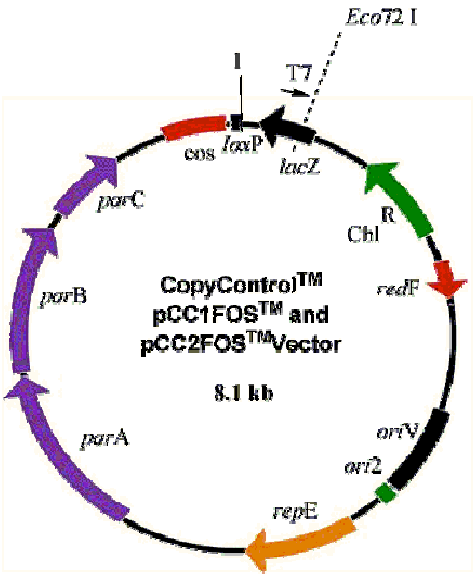


Figure 5-3 pCC2FOS™ Vector Map[252]

5.5.1 Screening of fosmid library

Single fosmid clones were picked up and used to inoculate 96 well plates containing LB medium with 12.5 µg/mL chloramphenicol (vetor is chloramphenicol resistant). Plates were covered up and sealed with tape around leaving a small gap for air circulation. These plates were put on shaker at 37ºC for overnight.

After incubation 16 µL from each well of 96-well plate was collected into a 1.5mL eppendorf tube and fosmid purification was performed following the protocol mentioned before. PCR reaction was carried out using specific primers followed by analysis on agarose gel electrophoresis using 1% agarose and ethidium bromide. Gel was analysed under UV. When there was a positive hit from the DNA sample representing the whole 96-wells, next step was to do one PCR reaction for a row of 12 wells and in the next stage on 12 individual wells to locate a single positive clone.

Positive band on agarose gel was cut and DNA was purified from the gel using 'gel purification kit' from Qiagen Ltd, U.K. and sent for sequencing to GATC biotech, Germany.

5.6 Marfey's Method

Marfey's method[241] was used to analyse the absolute stereochemistry of the constituent amino acids of Ga-foroxymithine. Ga-foroxymithine (500 µg) was dissolved in 400 µl of 6N HCl and heated at 110°C for 24 hrs. Mixture was lyophilized (freeze dried). Residue dissolved in 10 µL of H_2O followed by addition of 20 µL of 1 M $NaHCO_3$ and 170 µL of 1% Marfey's Reagent (Nα-(2,4-dinitro-5-fluorophenyl)-L-alaninamide, Sigma-Aldrich) acetone solution and heated at 37°C for 1 h. Reaction was quenched with 20 µL of 1 M HCl and the mixture lyphilized (freeze dried). The dried products were dissolved in 1:1 water:acetonitrile 0.1% TFA solution to a final volume of 400 µL.

Marfey's derivatisation of standard amino acids was achieved by mixing 50µL of 50mM (aq) amino acid, 100µL of 1% Marfey's reagent/acetone solution and 20µL of 1M $NaHCO_3$. This mixture was heated at 37°C for 1h and lyophilized. Lyophilized product was finally dissolved in 1:1 water:acetonitrile 0.1% TFA solution to a final volume of 400µL

Samples were then analysed on LC-MS using C18 HPLC RP column (250 x 4.6mm, 5 µm) on Agilent equipped with a Shimadzu SPD-10A UV-Vis detector. The separation was performed using a linear gradient of 0-52.5% buffer B (10 mM ammonium formate, 1% methanol, 60% acetonitrile, pH 5.2) in buffer A (10 mM ammonium formate, 1% methanol, 5% acetonitrile, pH 5.2) over 45 minutes by injecting 5 µL and UV detector set at wavelength 340 nm and flow rate of 1 mL/min.

5.7 LC-MS analysis

5.8.2.1 Sample analysis for fish toxins

Ethyl acetate filtrates and extracts were analysed by LC-MS using an Agilent 1100 HPLC connected via a splitter to a Bruker HCT+ mass spectrometer on an Agilent Eclipse XDB-C18 column (particle size 5 µm, size 4.6 x 150 mm) using electrospray ionisation.

Following method and solvents were used for LCMS analysis

Table 5.7 LC-MS method for fish toxins analysis

Time (minutes)	A (%)	B (%)
0	90	10
5	80	20
30	0	100
35	0	100
37	90	10
50	90	10
Flow rate 1 mL/min, UV 210nm A= water; B= Methanol		

5.8 NMR spectroscopy

1D and 2D NMR spectra of Ga-foroxymithine were recorded at 700 MHz Bruker AV700 spectrometer with a cryoprobe. Number of scans for each NMR experiment were; 128 for ^{1}H, HMBC and NOESY spectra, 32 for HSQC and 16 for COSY spectra. Other NMR spectra were recorded at 300 or 400 MHz using Bruker DPX300, or DPX400 spectrometers, respectively. Chemical shifts are quoted in ppm with reference to the residual solvent peak. The data in parentheses is in the following order: 1-multiplicity: s(singlet); d (doublet); t(triplet); q(quartet); m(multiplet), 2-number of equivalent protons, 3- coupling constant (J) in Hz, 4-assignment.

5.9 PCR

The following PCR reaction condition were used to amplify desired DNA fragments using primers listed in the table 4.5.

Table 5.8 List of reagents for PCR reaction

Reagents	Quantity
PCR buffer with $MgCl_2$	5 µL
DMSO	2.5 µL
Primers	2 x 1 µL
Taq polymerase	0.5 µL
DNA (cell culture)	1 µL
dNTPs	1 µL
MilliQ H_2O	38 µL
Total	**50 µL**

5.9.1 Standard PCR Method

A general method for PCR reaction set was as follows

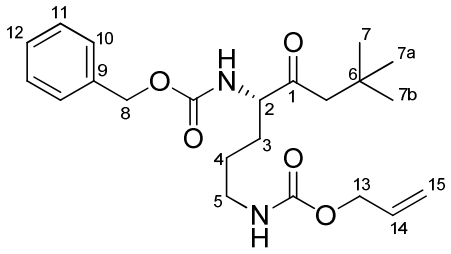
Table 5.9 Standard PCR method

		Temperature	Time
1	Denaturation	94°C	2 min
2	Denaturation	94°C	45 sec
3	Primer annealing	55-68°C	45 sec
4	Extension	72°C	90-240 sec
5	Go to step 2 for 35 cycles		
6	Final extension	72°C	15 min

5.10 Chemical synthesis of N5-hydroxyornithine

Both D and L stereoisomers of 5-N-OH-ornithine were chemically synthesised from commercially available D- and L-ornithine. The scheme is give below.

5.10.1 N2-(benzyloxycarbonyl)-N5-allyloxycarbonyl)ornithine t-butyl ester (3)



To the solution of L/D Ornithine HCl salt (1g, 5.93 mM) in NaOH (0.5M ,12 mL) was added $CuSO_4$. $5H_2O$ (0.74g, 2.97 mM) and left on stirring. After 15 minutes anhydrous K_2CO_3 (0.82g, 5.93 mM) was added followed by the addition of allyl chloroformate (800 µL 7.41 mM) dropwise. The mixture was left stirring for 3h under argon. It was then filtered (using Buchner funnel) and dried under vacuum. To this EDTA (0.87g, 2.97 mM) and NaOH (0.24g, 5.93 mM) solubilised in dH_2O (24 mL) were added. Resulting light blue mixture was then stirred under reflux (2h, 95°C). Dark coloured resulting solution with colourless solid was brought to room temperature. To this K_2CO_3 (0.41g, 2.97 mM) and benzyl oxycarbonyl chloride (cbz-

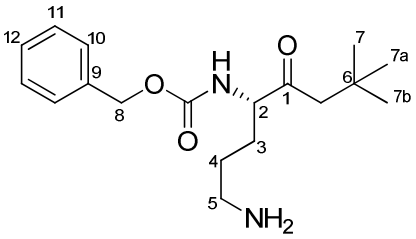
Cl, 460µL, 3.26 mM)) was added and left stirring overnight. Overnight reaction was quenched with pH 2 buffer (0.75M Na_2SO_4, 0.25M H_2SO_4, pH2). The resulting compound was extracted using DCM and dried under vacuum leaving behind an oily substance. After weighing the dried compound, it was redissolved in DCM (150 mL) and while still working under argon, trichloro acetimidate (1.9 mL, 10.5 mM) was slowly added to the reaction flask set over oil bath at 45°C. When the flask cooled down a bit, temperature of oil bath was increased to 55°C and a reflux condenser was set up over the reaction flask. TLC indicated the reaction was complete after 5 hours. Reaction mixture was washed with NaOH (aq., 1N, 100 mL). The organic layer was dried over anhydrous sodium sulphate and evaporated to dryness leaving behind a colourless oily residue. It was then purified by flash column chromatography (silica, 30% ethyl acetate/petroleum ether) yielding a colourless oil.

Yield: 1.27g (89%)

^{1}H-NMR ($CDCl_3$, 400MHz) ∂: 1.46 (m, 9H H-7/H-7a/H-7b), 1.56 (m, 4H, H-3/H-4), 3.2 (m, 2H, H-5), 4.25 (q, 1H, 7 Hz, H-2), 4.56 (d, 2H, H-13), 5.10 (s, 2H, H-8), 5.18 - 5.28 (m, 2H, H-15), 5.91 (m, 1H, H-14), 7.25 (s, H-8), 7.32 (m, 5H, H-10 - H12).

HRMS: *m/z* calculated for $C_{21}H_{30}N_2O_6$: 429.1996 $[M+Na]^+$. Found 429.1997.

5.10.2 N2-(benzyloxycarbonyl) ornithine t-butyl ester (4)



To a solution of N2-(benzyloxycarbonyl)-N5-allyloxycarbonyl)ornithine t-butyl ester in THF (50 mL) $NaBH_4$ (0.84g, 23.4 mM) was added and the mixture was left stirring for 10 minutes. To this mixture tetrakis triphenylphosphine palladium (0.18 g, 0.16 mM) was added resulting in a yellow solution which was stirred at room

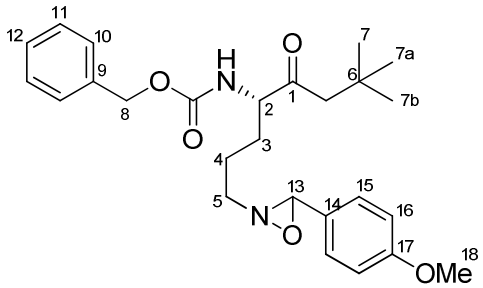
temperature and monitored by TLC. TLC indicated the completion of reaction after ~ 3h. The reaction was quenched with $NaHCO_3$ (sat) resulting in a whitish lumpy material. The resulting mixture was extracted using equal volume of DCM (3 times). It was then dried over anhydrous sodium sulphate and evaporated to leave an oily residue. The residue was purified by flash column chromatography (silica, 2%, 5% and 10% methanol (sat with NH_3/DCM)) and was evaporated to dryness yielding a colourless oil.

Yield: 0.84g (84%)

^{1}H-NMR ($CDCl_3$, 400MHz) ∂: 1.46 (s, 9H, H-7/H-7a/H-7b), 1.67 (m, 4H, H-3/H-4), 2.73 (m, 2H, H-5), 4.26 (m, 1H, H-2), 5.10 (s, 1H, H-8), 7.32 (m, 5H, H-10 - H12).

HRMS: *m/z* calculated for $C_{17}H_{26}N_2O_4$: 323.1965 $[M+H]^+$. Found 323.1961.

5.10.3 Oxaziridine (6)



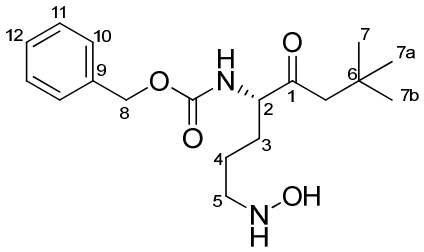
To a solution of N2-(benzyloxycarbonyl) ornithine t-butyl ester in methanol (2.5 mL) was added p-anisaldehyde (331 µL, 2.74 mM) and molecular mesh (3 Å). It was left on stirrer under argon for 23 hours at room temperature. Reaction mixture was then filtered over celite and dried under vacuum resulting in orange brown oily substance. Still working under argon, the resulting compound was dissolved in methanol (4 mL) and the reaction flask was put on ice. M-chloroperbenzoic acid (0.46 g, 2.67 mM) was dissolved in methanol (1 mL) and added into the reaction flask on ice under argon resulting in light yellow mixture. TLC indicated the completion of reaction. Compound was purified by flash column chromatography (silica, 2%, 3% and 4% ethyl acetate/petroleum ether) to leave the desired compound as viscous oil.

Yield: 1.07g (94%)

^{1}H-NMR ($CDCl_3$, 400MHz) ∂: 1.45 (s, 9H, H-7/H-7a/H-7b), 1.79 (m, 4H, H-3/H-4) 2.65, 3.07 (m, 2H, H-5), 3.81 (m, 3H, H-18), 4.38 (m, 1H, H-2), 4.46 (s, 1H, H-13), 5.10 (s, 2H, H-8), 7.34 (m,5H, H-10 - H12).

HRMS: *m/z* calculated for $C_{25}H_{32}N_2O_5$: 441.2384 $[M+H]^+$. Found 441.2385.

5.10.4 N2-(benzyloxycarbonyl)-N5-hydroxyornithine t-butyl ester (7)



To a solution of the above oxaziridine compound (200 mg) in dry methanol (4.5 mL) was added hydroxylamine hydrochloride, $HONH_2$ (0.046 mg, 0.66 mM) and left on stirring at room temperature. The reaction was monitored by TLC. On completion of the reaction (about 4h) it was evaporated under vacuum yielding an oily residue, which was purified by flash column chromatography (silica, 2%, 5% and 10% methanol/DCM) to leave the desired compound as an oily residue.

Yield: 0.122g (82%)

^{1}H-NMR ($CDCl_3$, 400MHz) ∂: 1.45 (s, 9H, H-7/H-7a/H-7b), 1.90 (m, 4H, H-3/H-4), 3.25 (m, 2H, H-5), 4.27 (m, 1H, H-2), 5.09 (s, 2H, H-8), 7.33 (m, 5H, H-10 - H-12).

HRMS: *m/z* calculated for $C_{17}H_{26}N_2O_5$: 339.1914 $[M+H]^+$. Found 339.1908.

CD spectra (Figure 5-4, 5-5) of N2-(benzyloxycarbonyl)-N5-hydroxyornithine t-butyl ester containing L- and D-ornithine as starting material were collected, which confirmed the synthesis of two different isomers.

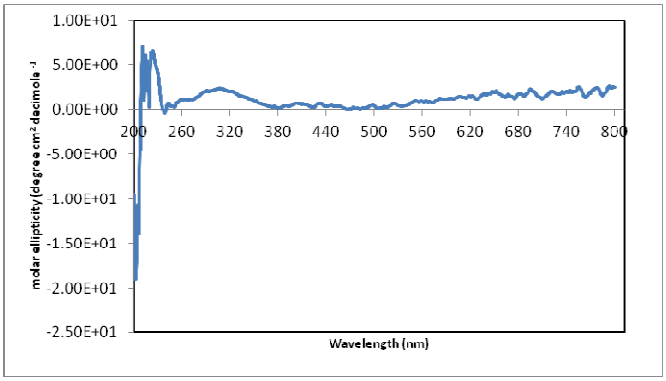


Figure 5-4 Circular dichroism spectrum of N2-(benzyloxycarbonyl)-N5-hydroxy-$_{L}$-ornithine t-butyl ester (7)

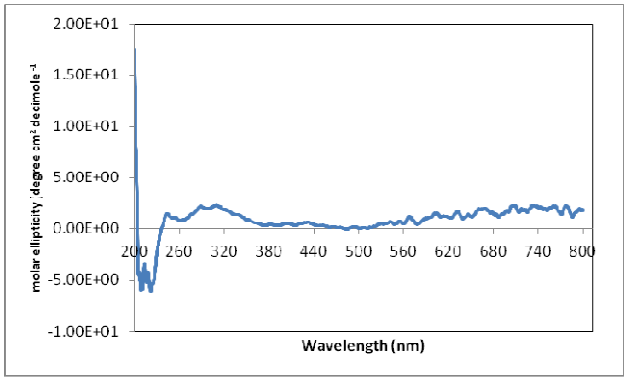
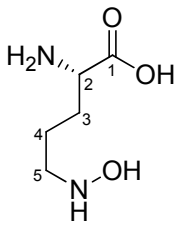


Figure 5-5 Circular dichroism spectrum of N2-(benzyloxycarbonyl)-N5-hydroxy-$_{D}$-ornithine t-butyl ester (7)

5.10.5 N5-hydroxyornithine (8)



To a solution of N2-(benzyloxycarbonyl)-N5-hydroxyornithine t-butyl ester (10 mg) in DCM (585 µL) was added 33% HBr (180 µL) solution in acetic acid and was left on stirring at room temperature. The reaction was monitored by TLC. On completion the mixture was diluted with chloroform and evaporated to dryness under vacuum. Residue was taken back in chloroform (x2) and toluene (x2) and evaporated to leave behind the desired final compound as a colourless viscous oil.

Yield: 0.008g (80 %)

^{1}H-NMR ($CDCl_3$, 400MHz) ∂: 1.76 (m, 4H, H-3/H-4), 3.6 (m, 1H, H-2), 4.26 (m, 2H, H-5)

HRMS: *m/z* calculated for $C_5H_{12}N_2O_3$: 149.0921 $[M+H]^+$. Found 149.0919.

CHAPTER 6 - RERERENCES

1. R. K. Hart, and Oglesby, G.B., *Scan. Electron Micros.*, 1979, **3**, 355-361.
2. V. Roncero, Vincente, J.A., rodondo, E., Gazquez, A., and Duran, E., *Environ. Health Perspect*, 1990, **89**, 137-144.
3. H. E. David, *Environ. Health Perspect*, 1987, **71**, 47-58.
4. J. F. Skidmore, Tovell,P.W. A., *Wat. Res.*, 1972, **6**, 217-230.
5. G. Smart, *J. Fish Biol.*, 1976, **8**, 471-475.
6. A. G. Ellis, Smith, D.G., *J. Exp. Zool.*, 1983, **227**, 371-380.
7. R. M. Stagg, Shuttleworth, T.J., *J. Comp. Physiol.*, 1983, **149**, 83-90.
8. P. G. Daye, Garside, E. T., *Can. J. Zool.*, 1976, **54**, 2140-2155.
9. P. D. Abel and J. F. Skidmore, *Water Res*, 1975, **9**, 759-765.
10. C. H. Jagoe, Haines, T.A., *Trans. Am. Fish. Soc.*, 1983, **112**, 689-695.
11. G. Chevalier, Gauthier, L., and Moreau, g., *Can. J. Zool.*, 1985, **63**, 2062-2070.
12. C. L. Bolis, Cambria, A., and Fama, M., *Effects of acid stress on fish gills. In: Toxins, Drugs and Pollutants in Marine Animals*, Springer-Verlag, Berlin, Berlin, 1984.
13. S. P. Ram, and Tarun, K.B., *Vet. Arhiv*, 2002, **72**, 167-183.
14. J. Mallatt, *Can. J. Fish. Aquat. Sci.*, 1985, **42**, 630-648.
15. G. M. Hughes, Perry, S.F., and Brown, V.M., *Wat. Res.*, 1979, **13**, 665-679.
16. L. Karlsson-Norrgren, Runn, P., Haux, c., and Forlin, L., *J. Fish Biol.*, 1985, **27**, 81-95.
17. P. Y. Daost, Wobser, G., and Newstead, J.D., *Vet. Pathol.*, 1984, **21**, 93-101.
18. j. w. Lewis, Kirk, R. S., *Environ. Technology*, 1993, **14**, 577-585.
19. J. L. J. Parry, *Investigating links between bacterial exotoxins and unexplained fish kills: phase 4*, Environment Agency: Rio House, Waterside Drive, Axtec West, Almondsbury, BS32 4UD, Bristol, UK, 2005.
20. E. Deak, I. SzaboA, A. Kalmaczhelyi, Z. Gal, G. Barabas and A. Penyige, *Microbiology*, 1998, **144**, 2169-2177.
21. L. A. McCue, J. Kwak, M. J. Babcock and K. E. Kendrick, *Gene*, 1992, **115**, 173-179.
22. T. B. H. O. Yoshiro Okami, Biology of Actinomycetes '88., 1988.
23. M. F. Watanabe, K. Tsuji, Y. Watanabe, K.-I. Harada and M. Suzuki, *Nat. Toxins*, 1992, **1**, 48-53.
24. M. F. Watanabe, S. Oishi, K.-I. Harada, K. Matsuura, H. Kawai and M. Suzuki, *Toxicon*, 1988, **26**, 1017-1025.
25. F. Tencalla and D. Dietrich, *Toxicon*, 1997, **35**, 583-595.
26. R. J. Andersen, H. A. Luu, D. Z. X. Chen, C. F. B. Holmes, M. L. Kent, M. Le Blanc, F. J. R. M. Taylor and D. E. Williams, *Toxicon*, 1993, **31**, 1315-1323.
27. B. Johnson I., S., Sims, I., Conrad, A., Parr, W., Hedgecott, S. and Cartwright, N., *Technical invetigation of the Hungerford fish mortality.*, Report to the Environment Agency, No CO 4626, 1998.
28. E. A. News-release, *Rare natural phenomena killed fish in Kennet and Avon Canal and River Dun pollution incident*, News release, Environment Agency: Rio House, Waterside Drive, Axtec West, Almondsbury, , BS32 4UD, Bristol, UK, Bristol, 1998.
29. J. W. Lewis, N. J. Morley, J. Drinkall, B. J. Jamieson, R. Wright and J. D. Parry, *Ecotox. Environ. Safe.*, 2009, **72**, 173-181.

30. T. Kieser, Mervyn j. Bibb, Mark j. Buttner, Keith F. chter, David A. Hopwood, *Practical Streptomyces Gentics*, The John Inns Foundation, Norwich, 2000.
31. S. A. Waksman, *Bacteriol Rev.*, 1957, **21**, 1-29.
32. N. A. K. Nikov, *J. Bacteriol.*, 1960, **79**, 65-74.
33. K. F. Chater, *Philos. T. Roy. Soc. B.*, 2006, **361**, 761-768.
34. S. Omura, *Gene*, 1992, **115**, 141-149.
35. D. A. Hopwood, *Streptomyces in Nature and Medicine: The Antibiotic Markers*, Oxford University Press, New York, 2007.
36. I. G. M.P. Lechevalier, M., Williams, S.T. and Mordarski M., in *Actinomycetes in Biotechnology*, Academic Press, San Diego, Editon edn., 1988, pp. 327-358.
37. B. W. Kennedy, and Alcorn S.M., *Plant Diseases*, 1980, **64**, 674-676.
38. A. C. Thaysen and F. T. K. Pentelow, *Ann. Appl. Biol.*, 1936, **23**, 105-109.
39. N. N. Gerber and H. A. Lechevalier, *Appl. Environ. Microbiol.*, 1965, **13**, 935-938.
40. R. G. Buttery and J. A. Garibaldi, *J. Agri. Food Chem.*, 1976, **24**, 1246-1247.
41. P.-E. Persson, *Water Res.*, 1980, **14**, 1113-1118.
42. Y. Xie, J. He, J. Huang, J. Zhang and Z. Yu, *J. Agr.Food Chem.*, 2007, **55**, 6823-6828.
43. S. T. Williams, M. Shameemullah, E. T. Watson and C. I. Mayfield, *Soil Biol. Biochem.*, 1972, **4**, 215-225.
44. M. Kontro, U. Lignell, M. R. Hirvonen and A. Nevalainen, *Lett. Appl. Microbiol.*, 2005, **41**, 32-38.
45. T. G. P. a. D. Gottleib, *J Bacteriol*, 1948, **56**, 107.
46. J. C. Ensign, *Ann Rev Microbiol*, 1978, **32**, 185-219.
47. E. Kuster and S. T. Williams, *Nature*, 1964, **202**, 928-929.
48. M. P. A. D. Mythili B. , *Res. J. Agr. Sci.*, 2011, **2**, 104-106.
49. K. Goshi, T. Uchida, A. Lezhava, M. Yamasaki, K. Hiratsu, H. Shinkawa and H. Kinashi, *J. Bacteriol.*, 2002, **184**, 3411-3415.
50. D. A. Hopwood, *Streptomyces in Nature and Medicine-The Antibiotic Makers*, Oxford University Press, 2007.
51. P. E. Turpin, V. K. Dhir, K. A. Maycroft, C. Rowlands and E. M. H. Wellington, *FEMS Microbiol. Lett.*, 1992, **101**, 271-280.
52. O. J. Raatikainen, T. H. Päivinen and R. T. Tahvonen, *Pestic. Sci.*, 1994, **41**, 149-154.
53. E. C. Eckwall and J. L. Schottel, *J. Ind. Microbiol. Biot.*, 1997, **19**, 220-225.
54. Y. Chen, Y. Luo, J. Ju, E. Wendt-Pienkowski, S. R. Rajski and B. Shen, *J. Nat. Prod.*, 2008, **71**, 431-437.
55. H. Umezawa, Aoyagi, T., Ogawa, K., Obata, T., Iinuma, H., Naganawa, H., Hamada, M., Takeuchi, T., *J. Antibiot. (Tokyo)*, 1985, **38**, 1813-1815.
56. A. Paananen, R. Mikkola, T. Sareneva, S. Matikainen, M. Andersson, I. Julkunen, M. S. Salkinoja-Salonen and T. Timonen, *Infect. Immun.*, 2000, **68**, 165-169.
57. S. Horinouchi and T. Beppu, *Antonie van Leeuwenhoek*, 1993, **64**, 177-186.
58. K. F. Chater, S. Biró, K. J. Lee, T. Palmer and H. Schrempf, *FEMS Microbiol. Rev.*, **34**, 171-198.
59. K. Flardh and M. J. Buttner, *Nat. Rev. Microbiol.*, 2009, **7**, 36-49.
60. C. I. Mayfield, S.T. Williamsa, S.M. Ruddicka and H.L. Hatfield, *Soil Biol. Biochem.*, 1972, **4**, 79-86.

61. J. F. Mcgregor, *J Gen Microbiol*, 1954, **11**, 52-56.
62. C. I. Mayfield, S. T. Williams, S. M. Ruddick and H. L. Hatfield, *Soil Biol. Biochem.*, 1972, **4**, 79-86, IN73-IN76, 87-91.
63. R. Y. Morita, *Starvation and miniaturisation of heterotrophs, with special emphasis on maintenance of the starved viable state. In Fletcher, M. and Floodgate, G.D. (eds)* Academic Press, London, London, 1985.
64. J. Hunter, ed. *S.* l. cycle, Jim Hunter, Editon edn., 2007, pp. Image published with the following permission from the author"I Jim Hunter of the University of East Anglia hereby give Mansoor Ahmad permission to use the "*Streptomyces* Life Cycle" image in its entire, unaltered state (excluding resizing) as created by Jim Hunter (2007); for his PhD thesis."
65. A. D. Karagouni, Vionis, A.P., Baker, P.W., Welington, E.M.H., *Microbiol. Releases*, 1993, **2**, 47-51.
66. E. G. Triger, Polyanskaya, L.M., Kozhevin, P.A., Zvyagintsev, D.G., *Mikrobiol*, 1991, **60**, 461-465.
67. L. Sanchez and A. F. Brana, *Microbiology*, 1996, **142**, 1209-1220.
68. S. Swift, J. Allan Downie, N. A. Whitehead, A. M. L. Barnard, G. P. C. Salmond and P. Williams, in *Adv. Microb. Physiol.*, Academic Press, Editon edn., 2001, vol. Volume 45, pp. 199-270.
69. C. Jiang, Xu, L., *Appl. Environ. Microbiol.*, 1996, **62**, 249-253.
70. H. M. Kieser, Kieser, T., Hopwood, D.A., *J. Bacteriol.*, 1992, **174**, 5496-5507.
71. P. Leblond, Redenbach, M., Cullum, J., *J. Bacteriol.*, 1993, **175**, 3422-3429.
72. P. Leblond, Demuyter, P., Simonet, J-M, Decaris, B., *J. Bacteriol.*, 1991, **173**, 4229-4233.
73. A. Lezhava, T. Mizukami, T. Kajitani, D. Kameoka, M. Redenbach, H. Shinkawa, O. Nimi and H. Kinashi, *J. Bacteriol.*, 1995, **177**, 6492-6498.
74. S. D. Bentley, K. F. Chater, A. M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill and D. A. Hopwood, *Nature*, 2002, **417**, 141-147.
75. A. Correia, J. F. Martin and J. M. Castro, *Microbiology*, 1994, **140**, 2841-2847.
76. M. Redenbach, J. Scheel and U. Schmidt, *Antonie van Leeuwenhoek*, 2000, **78**, 227-235.
77. H. Yu, M. Yuan, W. Lu, J. Yang, S. Dai, Q. Li, Z. Yang, J. Dong, L. Sun, Z. Deng, W. Zhang, M. Chen, S. Ping, Y. Han, Y. Zhan, Y. Yan, Q. Jin and M. Lin, *J. Bacteriol.*, 2011, JB.05039-05011.
78. H. Ikeda, J. Ishikawa, A. Hanamoto, M. Shinose, H. Kikuchi, T. Shiba, Y. Sakaki, M. Hattori and S. Omura, *Nat. Biotech.*, 2003, **21**, 526-531.
79. Y. Ohnishi, Ishikawa, Jun, Hara, Hirofumi, Suzuki, Hirokazu, Ikenoya, Miwa, Ikeda, Haruo, Yamashita, Atsushi, Hattori, Masahira, Horinouchi, Sueharu, *J. Bacteriol.*, 2008, **190**, 4050-4060.
80. K. J. Grubbs, P. H. W. Biedermann, G. Suen, S. M. Adams, J. A. Moeller, J. L. Klassen, L. A. Goodwin, T. Woyke, A. C. Munk, D. Bruce, C. Detter, R. Tapia, C. S. Han and C. R. Currie, *J. Bacteriol.*, 2011, **193**, 2890-2891.

81. D. A. Hopwood, *Nat. Biotech.*, 2003, **21**, 505-506.
82. G. L. C. a. D. A. Hopwood, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 14555-14561.
83. S. Lautru, R. J. Deeth, L. M. Bailey and G. L. Challis, *Nat. Chem. Biol.*, 2005, **1**, 265-269.
84. G. L. Challis, *Microbiol.*, 2008, **154**, 1555-1569.
85. C. Corre and G. L. Challis, *Nat. Prod. Rep.*, 2009, **26**, 977-986.
86. M. Zerikly and G. L. Challis, *Chem. BioChem.*, 2009, **10**, 625-633.
87. S. ÅŒmura, H. Ikeda, J. Ishikawa, A. Hanamoto, C. Takahashi, M. Shinose, Y. Takahashi, H. Horikawa, H. Nakazawa, T. Osonoe, H. Kikuchi, T. Shiba, Y. Sakaki and M. Hattori, *P. Natl. Acad. Sci.*, 2001, **98**, 12215-12220.
88. R. W. Burg, B. M. Miller, E. E. Baker, J. Birnbaum, S. A. Currie, R. Hartman, Y.-L. Kong, R. L. Monaghan, G. Olson, I. Putter, J. B. Tunac, H. Wallick, E. O. Stapley, R. Oiwa and S. Omura, *Antimicrob. Agents Chemother.*, 1979, **15**, 361-367.
89. S. Omura, H. Ikeda, J. Ishikawa, A. Hanamoto, C. Takahashi, M. Shinose, Y. Takahashi, H. Horikawa, H. Nakazawa, T. Osonoe, H. Kikuchi, T. Shiba, Y. Sakaki and M. Hattori, *P. Natl. Acad. Sci.*, 2001, **98**, 12215-12220.
90. J. Claesen and M. J. Bibb, *J. Bacteriol.*, **193**, 2510-2516.
91. E. F. Dunne, W. J. Burman and M. L. Wilson, *Clin. Infect. Dis.*, 1998, **27**, 93-96.
92. W. Green, Jr, TE Adams, *Am. J. Clin. Pathol.*, 1964, **42**, 75-91.
93. R. Hay, Mahgoub ES, Leon G, al-Sogair S, Welsh O, *J. Med. Vet. Mycol.*, 1992, **30**, 41-49.
94. P. M. Kohn, M. Tager, M. L. Siegel and R. Ashe, *New Engl. J. Med.*, 1951, **245**, 640-644.
95. P. R. R. Clarke, G. B. R. Warnock, R. Blowers, and Margaret Wilkinson, *J. Neurol. Neurosurg. Psychiatry.*, 1964, **27**, 553-555.
96. D. H. LAMBERT, LORIA, R., *Int. J. Syst. Bacteriol.*, 1989, **39**, 387-392.
97. W. A. Millard, and Burr, B., *Ann. Appl. Biol.* , 1926, **13**, 580-644.
98. R. E. Gordon, and Horan, C, *J. Gen. Microbiol.*, 1968, **2**, 223-233.
99. R. R. King, C. H. Lawrence and L. A. Calhoun, *J. Agr. Food Chem.*, 1992, **40**, 834-837.
100. R. Loria, J. Kers and M. Joshi, *Annu. Revi. Phytopathol.*, 2006, **44**, 469-487.
101. M. A. Andersson, R. Mikkola, R. Kroppenstedt, F. A. Rainey, J. Peltola, J. Helin, K. Sivonen, and M. S. Salkinoja-Salonen, *Appl. Environ. Microbiol.*, 1998b, **64**, 4767-4773.
102. R. Rylander, *Indoor Air*, 1998, **4 (suppl)**, 59-65.
103. R. Rylander, *Environ. Health Perspect*, 1999, **107 (suppl)**, 501-503.
104. J. Peltola, Andersson, M. A., Haahtela, T., Mussalo-Rauhamaa, H., Rainey, F. A., Kroppenstedt, R. M., Samson, R. A., Salkinoja-Salonen, M. S., *Appl. Environ. Microbiol.*, 2001, **67**, 3269-3274.
105. M. R. Hirvonen, A. Nevalainen, N. Makkonen, J. Monkkonen, and K. Savolainen, *Environ. Toxicol. Pharmacol.*, 1997, **3**, 57-63.
106. P. Skov, L.N. Petersen, and P. Wolkoff, In Indoor Air 99: Proceedings of the 8th Conference on Indoor Air Quality and Climate, London, UK, 1999.
107. M. A. Andersson, M. Nikulin, U. Koljalg, M.C. Andersson, F.A. Rainey, K. Reijula, E.L. Hintikka, and M. Salkinoja-Salonen, *Appl. Environ. Microbiol.*, 1997, **63**, 387-397.

108. S. A. Waksman, H. C. Reilly and D. A. Harris, *J. Bacteriol.*, 1948, **56**, 259-269.
109. S. A. Waksman and R. E. Curtis, *Soil Sci.*, 1918, **6**, 309-320.
110. Nobelprize.org, Editon edn., 1952.
111. S. A. Waksman, H. C. Reilly and D. B. Johnstone, *J. Bacteriol.*, 1946, **52**, 393-397.
112. S. A. Waksman, A. Schatz and D. M. Reynolds, *Ann.NY Acad. Sci.*, 1946, **48**, 73-86.
113. J. D. Adcock and R. A. Hettig, *Arch. Intern. Med.*, 1946, **77**, 179-195.
114. T. B. Yoshiro Okami, Hiroshi Ogawara, *Biology Of Actinomycetes 88: Proceedings Of Seventh International Symposium On Biology Of Actinomycetes*, Intl Specialized Book Service Inc, 1989.
115. A. F. Demain, *In History of Modern Biotechnology*, Fichter, A. (Springer, BErlin), 2000.
116. F. W. Hinshaw HC, Pfuetze kh.. *Miss Valley Med J*, 1947, **69**, 160-166.
117. L. A. Kominek, *Antimicrob. Agents Chemother.*, 1975, **7**, 856-860.
118. L. A. Kominek, *Antimicrob. Agents Chemother.*, 1975, **7**, 861-863.
119. T. G. Obrig, W. J. Culp, W. L. McKeehan and B. Hardesty, *J. Biol. Chem.*, 1971, **246**, 174-181.
120. A. B. Campelo and J. A. Gil, *Microbiology*, 2002, **148**, 51-59.
121. Gil and D. Campelo, *Appl. Microbiol. Biotechnol.*, 2003, **60**, 633-642.
122. G. Bringmann, G. Lang, K. Maksimenka, A. Hamm, T. A. M. Gulder, A. Dieter, A. T. Bull, J. E. M. Stach, N. Kocher, W. E. G. Muller and H. P. Fiedler, *Phytochemistry*, 2005, **66**, 1366-1373.
123. C. Bruntner, T. Binder, W. Pathom-aree, M. Goodfellow, A. T. Bull, O. Potterat, C. Puder, S. Horer, A. Schmid, W. Bolek, K. Wagner, G. Mihm and H.-P. Fiedler, *J. Antibiot.*, 2005, **58**, 346-349.
124. W. H. Kim, J. H. Jung and E. Lee, *J. Org. Chem.*, 2005, **70**, 8190-8192.
125. R. M. Heisey, J. Huang, S. K. Mishra, J. E. Keller, J. R. Miller, A. R. Putnam and T. D. J. D'Silva, *J. Agr. Food Chem.*, 1988, **36**, 1283-1286.
126. Y. Hosoya, H. Adachi, H. Nakamura, Y. Nishimura, H. Naganawa, Y. Okami and T. Takeuchi, *Tetrahedron Lett.*, 1996, **37**, 9227-9228.
127. J. A. Wendt, P. J. Gauvreau and R. D. Bach, *J. Am. Chem. Soc.*, 1994, **116**, 9921-9926.
128. N. Menéndez, M. Nur-e-Alam, A. F. Braña, J. Rohr, J. A. Salas and C. Méndez, *Chem. Biol.*, 2004, **11**, 21-32.
129. R. J. Walczak, A. J. Woo, W. R. Strohl and N. D. Priestley, *FEMS Microbiol. Lett.*, 2000, **183**, 171-175.
130. J. F. Martin, *J. Bacteriol.*, 2004, **186**, 5197-5201.
131. A. Lounès, A. Lebrihi, C. Benslimane, G. Lefebvre and P. Germain, *Appl. Microbiol. Biotechnol.*, 1996, **45**, 204-211.
132. J. Martín, in *Advances in Biochemical Engineering, Volume 6*, eds. T. Ghose, A. Fiechter and N. Blakebrough, Springer Berlin / Heidelberg, Editon edn., 1977, vol. 6, pp. 105-127.
133. K. E. Kendrick and J. C. Ensign, *J. Bacteriol.*, 1983, **155**, 357-366.
134. K. P. S. G. Pulverer, international symposium on Actinomycete biology, Cologne, 1981.
135. J. F. Martin and A. L. Demain, *Biochem. Bioph. Res. Co.*, 1976, **71**, 1103-1109.

136. U. Gräfe, M. Roth, A. Christner and E. J. Bormann, *Z. Allg. Mikrobiol.*, 1981, **21**, 633-642.
137. S. Horinouchi and T. Beppu, *Crit. Rev. Biotechnol.*, 1990, **10**, 191-204.
138. F. Microbiologica, *Folia Microbiol.*, 1989, **34**, 353-447.
139. U. GRafe, G. Reinhardt, D. Krebs, I. Eritt and W. F. Fleck, *J. Gen. Microbiol.*, 1984, **130**, 1237-1245.
140. S. Horinouchi and T. Beppu, *Annual Review of Microbiology*, 1992, **46**, 377-392.
141. S. Horinouchi and T. Beppu, *Gene*, 1992, **115**, 167-172.
142. J.-y. Kato, N. Funa, H. Watanabe, Y. Ohnishi and S. Horinouchi, *P. Natl .Acad. Sci.*, 2007, **104**, 2378-2383.
143. S. Horinouchi and T. Beppu, *Mol. Microbiol.*, 1994, **12**, 859-864.
144. M. T. Lezhava A, Kajitani T, Kameoka D, Redenbach M, Shinkawa H, Nimi O, Kinashi H., *J. Bacteriol.*, 1995, **177**, 6492-6498.
145. S. Horinouchi and T. Beppu, *Annu. Rev. Microbiol.*, 1992, **46**, 377-392.
146. A. L. Khokhlov AS, Tovarova II, Kleiner EM, Kovalenko IV, Krasilnikova OI, Kornitskaya EY, Pliner SA, *Z Allg Mikrobiol*, 1973, **13**, 647-655.
147. F. Y. Ohnishi Y, Higashi T, Chun HK, Furihata K, Sakuda S, Horinouchi S, *J Antibiot (Tokyo).* 2004, **57**, 218-223.
148. S.-U. Choi, C.-K. Lee, Y.-I. Hwang, H. Kinosita and T. Nihira, *Arch. Microbiol.*, 2003, **180**, 303-307.
149. S. Horinouchi, *Front. Biosci.*, 2002, **7**, d2045-2057.
150. Y. Ohnishi, S. Kameyama, H. Onaka and S. Horinouchi, *Mol. Microbiol.*, 1999, **34**, 102-111.
151. B. Hong, S. Phornphisutthimas, E. Tilley, S. Baumberg and K. McDowall, *Biotechnol. Lett.*, 2007, **29**, 57-64.
152. G. P. van Wezel and K. J. McDowall, *Nat. Prod. Rep.*, 2011.
153. T. Beppu, *Gene*, 1992, **115**, 159-165.
154. K. Miyake, T. Kuzuyama, S. Horinouchi and T. Beppu, *J. Bacteriol.*, 1990, **172**, 3003-3008.
155. H. Onaka, M. Sugiyama and S. Horinouchi, *J. Bacteriol.*, 1997, **179**, 2748-2752.
156. R. F. A. Hubert Lechevalier, Charles T. Corke, Conrad M. Haenseler and Selman A. Waksman, *Mycologia*, 1953, **45**, 155-171.
157. A. L. Demain and E. Inamine, *Bacteriologic. Rev.*, 1970, **34**, 1-19.
158. J. B. Walker and M. S. Walker, *Biochemistry*, 1969, **8**, 763-770.
159. C. Walsh, *Antibiotics: Actions, origins, resistance*, ASM Press, Washington, DC, 2003.
160. S. A. Waksman and A. Schatz, *J. Am. Pharm. Assoc.*, 1945, **34**, 273-291.
161. T. Higashi, Y. Iwasaki, Y. Ohnishi and S. Horinouchi, *J. Bacteriol.*, 2007, **189**, 3515-3524.
162. O. Y. Lee HS, Horinouchi S., *J. Mol. Microbiol. Biotechnol.*, 2001, **3**, 95-101.
163. H. Krügel, P. Krubasik, K. Weber, H. P. Saluz and G. Sandmann, *BBA Mol. Cell Biol. L.*, 1999, **1439**, 57-64.
164. N. Funa, M. Funabashi, Y. Ohnishi and S. Horinouchi, *J. Bacteriol.*, 2005, **187**, 8149-8155.
165. M. Komatsu, M. Tsuda, S. ÅŒmura, H. Oikawa and H. Ikeda, *P. Natl. Acad.Sciences*, 2008, **105**, 7422-7427.
166. J. Jiang, X. He and D. E. Cane, *Nat. Chem. Biol.*, 2007, **3**, 711-715.

167. P. V. Zimba, L. Khoo, P. S. Gaunt, S. Brittain and W. W. Carmichael, *J. Fish Dis.*, 2001, **24**, 41-47.
168. M. S. Von Wittenau and H. Els, *J. Am. Chem. Soc.*, 1961, **83**, 4678-4680.
169. U. Hornemann, L. H. Hurley, M. K. Speedie and H. G. Floss, *J. Am. Chem. Soc.*, 1971, **93**, 3028-3035.
170. N. J. M. J. W. Lewis, J. Drinkall, J.J. Jamieson, R Wright, J.D. Parry, *Ecotox. Environ. Safe.*, 2009, **72**, 173-181.
171. E. Kim, Kim, H., Hong, S. P.,Kang, K. H., Kho, Y. H., and Park, Y. H., *Gene*, 1993, **132**, 21-31.
172. W. Weisburg, g., Barns, S. M., Pelletier, D. A., and Lane, D. J., *J. Bacteriol.*, 1991, **173**, 697-703.
173. S. Cheenpracha, R. P. Borris, T. T. Tran, J. M. Jee, H. F. Seow, H.-Y. Cheah, C. C. Ho and L. C. Chang, *J. Brazil. Chem. Soc.*, 2011, **22**, 223-229.
174. Y. Zhuang, X. Teng, Y. Wang, P. Liu, G. Li and W. Zhu, *Organic Letters*, **13**, 1130-1133.
175. M. O. Lozinskii, T. N. Kudrya, S. V. Bonadyk and P. S. Pel'kis, *Chemistry of Heterocyclic Compounds*, 1977, **13**, 433-436.
176. E. K. Dolence and M. J. Miller, *J. Org. Chem.*, 1991, **56**, 492-499.
177. T. Osono, Y. Oka, S. Watanabe, Y. Numazaki, K. Moriyama, H. Ishida, K. Suzaki, Y. Okami and H. Umezawa, *J. Antibiot.* , 1967, **20**, 174–180.
178. H. Eiki, H. Gushima, T. Saito, H. Ishida, Y. Oka and T. Osono, *J. Ferment. Technol.*, 1988, **66**, 559-565.
179. I. Maezawa, Hori T, Kinumaki, A., Suzuki, M., *j. Antibiot. (Tokyo)*, 1977, **26**, 771-775.
180. G. L. Challis and J. Ravel, *FEMS Microbiol.Lett.*, 2000, **187**, 111-114.
181. *Japan Pat.*, 1985.
182. D. S. Gottlieb, P. E., *Antibiotics. I. Mechanism of Action*, Springer-Verlag, Berlin, Heidelberg, New York., 1967.
183. H. Kleinkauf, and Gevers, W., Cold Spring Harb Symp Quant Biol, New York, 1969.
184. D. Schwarzer, Finking, R., and Marahiel, M.A., *Nat. Prod. Rep.*, 2003, **20**, 275-287.
185. F. Lipmann, *Adv. Microb. Physiol.*, 1980, **21**, 227-266.
186. M. A. Marahiel, *FEBS J.*, 1992, **307**, 40-43.
187. D. J. Smith, A. J. Earl and G. Turner, *EMBO J*, 1990, **9**, 2743-2750.
188. H. van Liempt, H. von Döhren and H. Kleinkauf, *J. Biol. Chem.*, 1989, **264**, 3680-3684.
189. R. Finking and M. A. Marahiel, *Ann. Rev. Microbiol.*, 2004, **58**, 453-488.
190. A. A. van Wageningen, P. N. Kirkpatrick, D. H. Williams, B. R. Harris, J. K. Kershaw, N. J. Lennard, M. Jones, S. J. M. Jones and P. J. Solenberg, *Chem. Biol.*, 1998, **5**, 155-162.
191. D. Bischoff, B. Bister, M. Bertazzo, V. Pfeifer, E. Stegmann, G. J. Nicholson, S. Keller, S. Pelzer, W. Wohlleben and R. D. Süssmuth, *ChemBioChem.*, 2005, **6**, 267-272.
192. B. K. Hubbard and C. T. Walsh, *Angew. Chem. Int. Ed.*, 2003, **42**, 730-765.
193. D. Konz, A. Klens, K. Schörgendorfer and M. A. Marahiel, *Chem. Biol.*, 1997, **4**, 927-937.
194. N. Kessler, H. Schuhmann, S. Morneweg, U. Linne and M. A. Marahiel, *J. Biol. Chem.*, 2004, **279**, 7413-7419.

195. G. Weber, K. Schörgendorfer, E. Schneider-Scherzer and E. Leitner, *Curr. Genetics*, 1994, **26**, 120-125.
196. L. E. N. Quadri, J. Sello, T. A. Keating, P. H. Weinreb and C. T. Walsh, *Chem. Biol.*, 1998, **5**, 631-645.
197. A. M. Gehring, K. A. Bradley and C. T. Walsh, *Biochemistry*, 1997, **36**, 8495-8503.
198. G. L. Challis, J. Ravel and C. A. Townsend, *Chem. Biol.*, 2000, **7**, 211-224.
199. M. A. Marahiel, T. Stachelhaus and H. D. Mootz, *Chem. Rev.*, 1997, **97**, 2651-2674.
200. T. Stachelhaus and M. A. Marahiel, *FEMS Microbiol. Lett.*, 1995, **125**, 3-14.
201. A. Tanovic, S. A. Samel, L.-O. Essen and M. A. Marahiel, *Science*, 2008, **321**, 659-663.
202. M. Strieker, A. Tanovic and M. A. Marahiel, *Curr. Opin. Struc. Biol.*, 2010, **20**, 234-240.
203. T. Stachelhaus, H. D. Mootz and M. A. Marahiel, *Chem. Biol.*, 1999, **6**, 493-505.
204. S. Lautru and G. L. Challis, *Microbiol.*, 2004, **150**, 1629-1636.
205. F. Barona-Gómez, U. Wong, A. E. Giannakopulos, P. J. Derrick and G. L. Challis, *J. Am. Chem. Soc.*, 2004, **126**, 16282-16283.
206. E. Conti, T. Stachelhaus, M. A. Marahiel and P. Brick, *EMBO J.*, 1997, **16**, 4174-4183.
207. K. Eppelmann, T. Stachelhaus and M. A. Marahiel, *Biochemistry*, 2002, **41**, 9718-9726.
208. J. J. May, N. Kessler, M. A. Marahiel and M. T. Stubbs, *Proc. Natl. Acad. Sci.*, 2002, **99**, 12120-12125.
209. T. V. Lee, L. J. Johnson, R. D. Johnson, A. Koulman, G. A. Lane, J. S. Lott and V. L. Arcus, *J. Biol. Chem.*, 2010, **285**, 2415-2427.
210. E. J. Dimise, P. F. Widboom and S. D. Bruner, *Proc. Natl. Acad. Sci. USA.*, 2008, **105**, 15311-15316.
211. L. Robbel, T. A. Knappe, U. Linne, X. Xie and M. A. Marahiel, *FEBS J.*, 2010, **277**, 663-676.
212. D. Schwarzer and M. A. Marahiel, *Naturwissenschaften*, 2001, **88**, 93-101.
213. T. Weber, R. Baumgartner, C. Renner, M. A. Marahiel and T. A. Holak, *Structure (London, England : 1993)*, 2000, **8**, 407-418.
214. S. A. Samel, G. Schoenafinger, T. A. Knappe, M. A. Marahiel and L.-O. Essen, *Structure*, 2007, **15**, 781-792.
215. P. J. Belshaw, C. T. Walsh and T. Stachelhaus, *Science*, 1999, **284**, 486-489.
216. T. A. Keating, C. G. Marshall, C. T. Walsh and A. E. Keating, *Nat. Struct. Mol. Biol.*, 2002, **9**, 522-526.
217. C. G. Marshall, M. D. Burkart, R. K. Meray and C. T. Walsh, *Biochemistry*, 2002, **41**, 8429-8437.
218. D. E. Ehmann, J. W. Trauger, T. Stachelhaus and C. T. Walsh, *Chem. Biol.*, 2000, **7**, 765-772.
219. U. Linne and M. A. Marahiel, *Biochem.*, 2000, **39**, 10439-10447.
220. E. D. Roche and C. T. Walsh, *Biochemistry*, 2003, **42**, 1334-1344.
221. T. Stachelhaus and C. T. Walsh, *Biochem.*, 2000, **39**, 5775-5787.
222. S. D. Uwe Linne, and, Mohamed A. Marahiel, *Biochemistry*, 2001, **40**, 15824-15834.
223. S. L. Clugston, S. A. Sieber, M. A. Marahiel and C. T. Walsh, *Biochemistry*, 2003, **42**, 12095-12104.

224. U. Linne, S. Doekel and M. A. Marahiel, *Biochemistry*, 2001, **40**, 15824-15834.
225. D. B. Stein, U. Linne, M. Hahn and M. A. Marahiel, *ChemBioChem.*, 2006, **7**, 1807-1814.
226. J. W. Trauger, R. M. Kohli, H. D. Mootz, M. A. Marahiel and C. T. Walsh, *Nature*, 2000, **407**, 215-218.
227. F. Kopp and M. A. Marahiel, *ChemInform*, 2007, **38**, 735-749.
228. L. Du and L. Lou, *Nat. Prod. Rep.*, 255-278.
229. F. Kopp, C. Mahlert, J. Grünewald and M. A. Marahiel, *J. Am. Chem. Soc.*, 2006, **128**, 16478-16479.
230. T. A. Keating, C. G. Marshall and C. T. Walsh, *Biochemistry*, 2000, **39**, 15522-15530.
231. S. D. Bruner, T. Weber, R. M. Kohli, D. Schwarzer, M. A. Marahiel, C. T. Walsh and M. T. Stubbs, *Structure*, 2002, **10**, 301-310.
232. S. A. Samel, B. Wagner, M. A. Marahiel and L.-O. Essen, *J. Mol. Biol.*, 2006, **359**, 876-889.
233. D. P. Frueh, H. Arthanari, A. Koglin, D. A. Vosburg, A. E. Bennett, C. T. Walsh and G. Wagner, *Nature*, 2008, **454**, 903-906.
234. S. M. Barry and G. L. Challis, *Curr. Opin. Chem. Biol.*, 2009, **13**, 205-215.
235. C. Corre and G. L. Challis, *Chem. Biol.*, 2007, **14**, 7-9.
236. V. Pohlmann and M. A. Marahiel, *Org. Biomol. Chem.*, 2008, **6**, 1843-1848.
237. C. B'Hymer, M. Montes-Bayon and J. A. Caruso, *J. Separation Sci.*, 2003, **26**, 7-19.
238. R. Bhushan and H. Brückner, *Amino Acids*, 2004, **27**, 231-247.
239. G. Gübitz, *Chromatographia*, 1990, **30**, 555-564.
240. C. Pettersson, *Trends Anal. Chem.*, 1988, **7**, 209-217.
241. P. Marfey, *Carlsberg Res. Comm.*, 1984, **49**, 591-596.
242. K. J. Ruterbories and D. Nurok, *Anal. Chem.*, 1987, **59**, 2735-2736.
243. Y. Nagata, T. Iida and M. Sakai, *J. Mol. Cat. B: Enzymatic*, 2001, **12**, 105-108.
244. H. Bruckner and C. Gah, *J. Chromatog. A*, 1991, **555**, 81-95.
245. A. Neuberger, in *Adv. Protein Chem.*, eds. M. L. Anson and T. E. John, Academic Press, Editon edn., 1948, vol. Volume 4, pp. 297-383.
246. Y. Nagata, K. Yamamoto and T. Shimojo, *Chromatog.*, 1992, **575**, 147-152.
247. G. Szokan, G. b. Mezˆ and F. Hudecz, *J. Chromatog. A*, 1988, **444**, 115-122.
248. P. Patel, University of Warwick, 2009.
249. H. F. Stephan, S.; Meyer, j. M.; Winkelmann, G.; Jung, G., *Liebigs Annalen der Chemie*, 1993, **1993**, 43-48.
250. O. Lazos, University of Cambridge 2009.
251. J. S. a. D. W. Russell, *Molecular Cloning - A Laboratory Manual*, 3rd edn., Cold Spring Harbor Laboratory Press, 2001.
252. EPICENTRE®-Biotechnologies, http://www.epibio.com/item.asp?id=385, Accessed 20 October 2010, 2010.